GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:06:39 ; Search time 15 Seconds
(without alignments)
1769.504 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1140	75.0	359	2	JC5382	hereditary hemochr
2	542.5	35.7	341	2	A57136	class I histocompa
3	523	34.4	361	1	HLRB	MHC class I histoc
4	523	34.4	361	2	I46858	MHC class I RLA pr
5	520	34.2	332	2	S06424	MHC class I histoc
6	517	34.0	365	2	I36961	MHC class I protei
7	516	33.9	361	2	B27638	MHC class I histoc
8	515	33.9	365	2	I83063	A11.2 - human
9	514	33.8	365	2	A47636	MHC class I histoc
10	514	33.8	365	2	I56039	HLA-A30.3 precurso
11	512	33.7	370	1	HLHUA3	MHC class I histoc
12	510	33.6	365	2	I38439	MHC class I histoc
13	509	33.5	365	2	I37542	MHC class I histoc
14	509	33.5	365	2	I38442	gene HLA-A-0205 pr
15	509	33.5	365	2	I61902	MHC class I histoc
16	508	33.4	365	2	I72170	MHC class I histoc
17	508	33.4	365	2	I38441	gene HLA-A-6802 pr
18	506	33.3	365	1	HLHUA2	MHC class I histoc
19	506	33.3	365	2	I37482	MHC class I histoc
20	506	33.3	365	2	I38519	MHC class I histoc
21	506	33.3	365	2	I84448	MHC class I histoc
22	505	33.2	365	2	I38610	MHC class I histoc
23	505	33.2	365	2	I37470	HLA-A*0210 - human
24	504	33.2	355	2	T28149	MHC class I histoc
25	504	33.2	364	2	S03535	class I histocompa
26	503	33.1	365	2	I37476	MHC class I histoc
27	503	33.1	365	2	I37478	MHC class I histoc
28	503	33.1	365	2	I38443	gene HLA-A-0203 pr
29	503	33.1	365	2	I61857	MHC HLA-A2.4a chai
30	502.5	33.1	341	2	JC5663	major histocompati
31	502	33.0	357	2	I36965	MHC class I protei
32	501.5	33.0	362	2	A45845	MHC class I histoc
33	501	33.0	365	2	I61856	MHC class I histoc
34	501	33.0	365	2	I54493	MHC class I histoc
35	500	32.9	273	1	HLHU69	MHC class I histoc
36	500	32.9	365	2	S77963	MHC class I histoc
37	500	32.9	365	2	S01171	class I histocompa
38	500	32.9	365	2	I54416	HLA-AW24 protein -
39	499	32.8	365	2	I37483	HLA-Aw34.2 antigen
40	498	32.8	273	1	HLHUAW	MHC class I histoc
41	498	32.8	360	2	A27638	MHC class I histoc
42	498	32.8	365	2	I72171	HLA-Aw33.1, HLA-Aw
43	497.5	32.7	339	2	I56071	MHC class I histoc
44	497	32.7	279	2	JX0353	zinc-alpha 2-glyco
45	497	32.7	362	2	I68724	MHC class I histoc

ALIGNMENTS

RESULT 1
JC5382
hereditary hemochromatosis protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C;Accession: JC5382
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Biochem. Biophys. Res. Commun. 230, 35-39, 1997
A;Title: Identification of a mouse homolog for the human hereditary haemochromatosis ca
A;Reference number: JC5382; MUID:97148566; PMID:9020055
A;Accession: JC5382
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-359 <HAS>
A;Cross-references: GB:U66849; NID:g1519484; PIDN:AAB07525.1; PID:g1519485
C;Comment: This protein plays a role in iron metabolism.
C;Genetics:
A;Gene: mr2
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-359/Product: hereditary haemochromatosis protein #status predicted <MAT>
F;30-117/Domain: alpha 1 #status predicted <ALF1>
F;118-217/Domain: alpha 2 #status predicted <ALF2>
F;218-309/Domain: alpha 3 #status predicted <ALF3>
F;314-340/Domain: transmembrane #status predicted <TRM>
F;341-359/Domain: intracellular #status predicted <INT>

Query Match 75.0%; Score 1140; DB 2; Length 359;
Best Local Similarity 72.2%; Pred. No. 2.3e-87;
Matches 203; Conservative 30; Mismatches 40; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |: |||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
               |||||||||:|| ||||||| |:|||         ||| |||:|||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:||||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310
```

```
RESULT 2
A57136
class I histocompatibility antigen related protein MR1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 23-Jul-1999
C;Accession: A57136
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Science 269, 693-695, 1995
A;Title: A gene outside the human MHC related to classical HLA class I genes.
A;Reference number: A57136; MUID:95350662; PMID:7624800
A;Accession: A57136
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-341 <HAS>
A;Cross-references: GB:U22963; NID:g940353; PIDN:AAC50174.1; PID:g940354
C;Genetics:
A;Gene: GDB:HLALS
A;Cross-references: GDB:683188; OMIM:600764
A;Map position: 1q25.3-1q25.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

  Query Match             35.7%;  Score 542.5;  DB 2;  Length 341;
  Best Local Similarity   39.5%;  Pred. No. 1.2e-37;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |    |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 3
HLRB
MHC class I histocompatibility antigen RLA alpha chain precursor (RL-5) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 22-Jun-1999
C;Accession: A02193
R;Tykocinski, M.L.; Marche, P.N.; Max, E.E.; Kindt, T.J.
J. Immunol. 133, 2261-2269, 1984
A;Title: Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expr
A;Reference number: A02193; MUID:84290724; PMID:6432910
A;Accession: A02193
A;Molecule type: mRNA
A;Residues: 1-361 <TYK>
A;Cross-references: GB:K02441; NID:g1293894; PIDN:AAA98729.1; PID:g165496
A;Note: the source of this protein is a T-lymphoid cell line (RL-5), which has been tran
C;Comment: In contrast to the many antigens expressed in mouse (K, D, and L) and human (
 MHC may therefore differ from the HLA and H-2 loci in having limited complexity.
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: class I histocompatibility antigen RLA alpha chain #status predicted <
F;25-307/Domain: extracellular #status predicted <EXT>
F;25-114/Domain: alpha-1 <EX1>
F;115-206/Domain: alpha-2 <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;308-329/Domain: transmembrane #status predicted <TMM>
F;330-361/Domain: intracellular #status predicted <INT>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

  Query Match             34.4%;  Score 523;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 5.3e-36;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||  |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 4
I46858
MHC class I RLA precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 21-Jan-2000
C;Accession: I46858
R;Marche, P.N.; Tykocinski, M.L.; Max, E.E.; Kindt, T.J.
Immunogenetics 21, 71-82, 1985
A;Title: Structure of a functional rabbit class I MHC gene: Similarity to human class I
A;Reference number: I46858; MUID:85103547; PMID:3917974
A;Accession: I46858
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-361 <MAR>
A;Cross-references: GB:K02819; NID:g165497; PIDN:AAA98730.1; PID:g165498
C;Genetics:
A;Introns: 25/1; 115/1; 207/1; 299/1; 337/1; 348/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             34.4%;  Score 523;  DB 2;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 5.3e-36;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||  |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 5
S06424
MHC class I histocompatibility antigen Ch25 alpha chain precursor - chimpanzee
N;Alternate names: MHC ChlA chain
C;Species: Pan troglodytes (chimpanzee)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999
C;Accession: S06424; I36959
R;Lawlor, D.A.; Ward, F.E.; Ennis, P.D.; Jackson, A.P.; Parham, P.
Nature 335, 268-271, 1988
A;Title: HLA-A and B polymorphisms predate the divergence of humans and chimpanzees.
A;Reference number: S06424; MUID:88319000; PMID:3412487
A;Accession: S06424
A;Molecule type: mRNA
A;Residues: 1-332 <LAW>
R;Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.
J. Immunol. 142, 3937-3950, 1989
A;Title: Diversity and diversification of HLA-A,B,C alleles.
A;Reference number: I36956; MUID:89235215; PMID:2715640
A;Accession: I36959
A;Molecule type: mRNA
A;Residues: 1-332 <RES>
A;Cross-references: GB:M24047; NID:g176818; PIDN:AAA35426.1; PID:g553155
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: glycoprotein; membrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-114/Domain: alpha-1 #status predicted <EX1>
F;115-206/Domain: alpha-2 #status predicted <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;307-331/Domain: transmembrane #status predicted <TMM>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

Query Match 34.2%; Score 520; DB 2; Length 332;
Best Local Similarity 40.1%; Pred. No. 8.5e-36;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::| |       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | ||||  |   |:: || |:. |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 6
I36961
MHC class I protein - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I36961
R;Lawlor, D.A.; Warren, E.; Ward, F.E.; Parham, P.
Immunol. Rev. 113, 147-185, 1990
A;Title: Comparison of class I MHC alleles in humans and apes.
A;Reference number: I36961; MUID:90201944; PMID:1690682
A;Accession: I36961
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30678; NID:g176822; PIDN:AAA87970.1; PID:g176823
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 34.0%; Score 517; DB 2; Length 365;
Best Local Similarity 39.7%; Pred. No. 1.7e-35;
Matches 110; Conservative 45; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::| |       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | ||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 7
B27638
MHC class I histocompatibility antigen alpha chain precursor (BL3-7) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 16-Feb-1997
C;Accession: B27638
R;Ennis, P.D.; Jackson, A.P.; Parham, P.
J. Immunol. 141, 642-651, 1988
A;Title: Molecular cloning of bovine class I MHC cDNA.
A;Reference number: A92826; MUID:88258075; PMID:3133413
A;Accession: B27638
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-361 <ENN>
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: MHC class I histocompatibility antigen, BoLA alpha chain (BL3-7) #stat
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 33.9%; Score 516; DB 2; Length 361;
Best Local Similarity 38.9%; Pred. No. 2e-35;
Matches 109; Conservative 50; Mismatches 113; Indels 8; Gaps 7;

```
Qy          2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISS 59
              |  |||| | : | |   ||   | |:|||||  || :| :  : | ||| ||:
Db         23 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 81

Qy         60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
              : | : ::  |     :| ||  |:   :| |:  || :| : ||::  |      |:|::|
Db         82 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 141

Qy        118 YDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
              |||:|::    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db        142 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 200

Qy        178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              | :  ||   |||| ::    |||| || :||: |:: | :: :    ::  |  :  |
Db        201 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 259

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              :||||:|  |  | || ||||||||:|:|  || :|| : ||
Db        260 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 299
```

RESULT 8

```
I83063
A11.2 - human
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C;Accession: I83063
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.;
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I83063
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16842; NID:g540517; PIDN:BAA04118.1; PID:g487911
C;Genetics:
A;Gene: A1102
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.9%;  Score 515;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 2.5e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGKPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 9
A47636
MHC class I histocompatibility antigen HLA-A11 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 28-Apr-1995 #text_change 23-Jul-1999
C;Accession: S03536; S03694; A47636; I60129
R;Mayer, W.E.; Jonker, M.; Klein, D.; Ivanyi, P.; van Seventer, G.; Klein, J.
EMBO J. 7, 2765-2774, 1988
A;Title: Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-spec
A;Reference number: S01171; MUID:89030641; PMID:2460344
A;Accession: S03536
A;Molecule type: mRNA
A;Residues: 1-365 <MAY>
A;Cross-references: EMBL:X13111; NID:g32138; PIDN:CAA31503.1; PID:g32139
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
A;Accession: S03694
A;Molecule type: mRNA
A;Residues: 1-42,'K',44-298 <MA2>
A;Cross-references: EMBL:X13112; NID:g32142; PIDN:CAA31504.1; PID:g32143
A;Note: this allele is designated A*1102 (formerly A11K, A11.2)
R;Cowan, E.P.; Jelachich, M.L.; Biddison, W.E.; Coligan, J.E.
Immunogenetics 25, 241-250, 1987
A;Title: DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows identification
A;Reference number: A47636; MUID:87192928; PMID:2437024
A;Accession: A47636
A;Molecule type: DNA
A;Residues: 26-365 <COW>
A;Cross-references: GB:M16007; GB:M16008; GB:M16009; GB:M16010; NID:g184130; PIDN:AAA654
A;Note: the authors translated the codon GAC for residue 89 as Ala, CCG for residue 104
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I60129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16841; NID:g540516; PIDN:BAA04117.1; PID:g487909
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: class I histocompatibility antigen alpha chain #status predicted <MAT
F;25-298/Domain: extracellular #status predicted <EXT>
F;220-285/Domain: immunoglobulin homology <IMM>
F;299-337/Domain: transmembrane #status predicted <TMM>
F;338-365/Domain: intracellular #status predicted <INT>

  Query Match             33.8%;  Score 514;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 3e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 10
I56039
HLA-A30.3 precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I56039
R;Kato, K.; Trapani, J.A.; Allopenna, J.; Dupont, B.; Yang, S.Y.
J. Immunol. 143, 3371-3378, 1989
A;Title: Molecular analysis of the serologically defined HLA-Aw19 antigens. A genetical
A;Reference number: I56039; MUID:90038496; PMID:2478623
A;Accession: I56039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30576; NID:g187646; PIDN:AAA59612.1; PID:g386878
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.8%;  Score 514;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 3e-35;
  Matches  109;  Conservative   48;  Mismatches  112;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
```

```
Db         26 SHSMRYFFTSVSRPGSGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|        ||  |:   :| |:  |||:|:: ||::  |      || :: |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYEQHAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  |: :||  :  | | |||||  |   |:: || |: |
Db        145 KDYIALNEDLRSWTAADMAAQITQRKWEAAR-WAEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| ||:||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 11
HLHUA3
MHC class I histocompatibility antigen HLA-A3 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 02-Sep-1997
C;Accession: A02192
R;Strachan, T.; Sodoyer, R.; Damotte, M.; Jordan, B.R.
EMBO J. 3, 887-894, 1984
A;Title: Complete nucleotide sequence of a functional class I HLA gene, HLA-A3: implicat
A;Reference number: A02192; MUID:84207948; PMID:6609814
A;Accession: A02192
A;Molecule type: DNA
A;Residues: 1-370 <STR>
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
A;Introns: 30/1; 120/1; 212/1; 304/1; 343/1; 354/1; 370/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-370/Product: class I histocompatibility antigen HLA-A3 alpha chain #status predicte
F;30-312/Domain: extracellular #status predicted <EXT>
F;30-119/Domain: alpha-1 <EX1>
F;120-211/Domain: alpha-2 <EX2>
F;225-290/Domain: immunoglobulin homology <IMM>
F;313-337/Domain: transmembrane #status predicted <TMM>
F;338-370/Domain: intracellular #status predicted <INT>
F;115/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;232-288/Disulfide bonds: #status predicted

  Query Match             33.7%;  Score 512;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 4.5e-35;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels   10;  Gaps    8;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:|||| ||:      : |
Db         31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|        ||  |:   :| |:  |||:|:: ||::  |      || :  |||
Db         90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              :|::    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db        150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy        180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db        208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||:| |  : || |||||||| |:| || :|| : ||
Db        267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 12
I38439
MHC class I histocompatibility antigen HLA-A*8001 precursor - human
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 21-Jan-2000
C;Accession: I59638; I38439
R;Domena, J.D.; Hildebrand, W.H.; Bias, W.B.; Parham, P.
Tissue Antigens 42, 156-159, 1993
A;Title: A sixth family of HLA-A alleles defined by HLA-A*8001.
A;Reference number: I59638; MUID:94112691; PMID:8284791
A;Accession: I59638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <DOM>
A;Cross-references: GB:L18898; NID:g306853; PIDN:AAA17012.1; PID:g306854
R;Balas, A.; Garcia-Sanchez, F.; Gomez-Reino, F.; Vicario, J.L.
Immunogenetics 39, 452, 1994
A;Title: Characterization of a new and highly distinguishable HLA-A allele in a Spanish
A;Reference number: I38439; MUID:94245293; PMID:8188325
A;Accession: I38439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <BAL>
A;Cross-references: EMBL:U03754; NID:g432407; PIDN:AAC04322.1; PID:g432408
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.6%;  Score 510;  DB 2;  Length 365;
  Best Local Similarity   38.3%;  Pred. No. 6.5e-35;
  Matches  106;  Conservative   53;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:|||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::::|        :  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||  : || | ||||| :|   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| ||:||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|:|||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 13
I37542
MHC class I histocompatibility antigen (allele A*0216) precursor - hu
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I37542; S49582
R;Barouch, D.; Krausa, P.; Bodmer, J.; Browning, M.J.; McMichael, A.J.
Immunogenetics 41, 388, 1995
A;Title: Identification of a novel HLA-A2 subtype, HLA-A*0216.
A;Reference number: I37542; MUID:95278976; PMID:7759139
A;Accession: I37542
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:Z46633; NID:g575248; PIDN:CAA86602.1; PID:g575249
```

```
A;Note: submitted to the EMBL Data Library, November 1994
C;Genetics:
A;Gene: hla-A
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.5%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 7.9e-35;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |    |    |   | |:|||||  || :| :  |:|:||| ||:     : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|      ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | ||||| :|   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGECVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
                :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 14
I38442
gene HLA-A-0205 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I38442
R;Holmes, N.; Ennis, P.; Wan, A.M.; Denney, D.W.; Parham, P.
J. Immunol. 139, 936-941, 1987
A;Title: Multiple genetic mechanisms have contributed to the generation of the HLA-A2/A2
A;Reference number: I38441; MUID:87252273; PMID:3496393
A;Accession: I38442
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:U03862; NID:g432436; PIDN:AAA03603.1; PID:g432437
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.5%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 7.9e-35;
  Matches  110;  Conservative   44;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :    |    |   | |:|||||  || :| :  |||:||| ||:     : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|      ||  |:   :| |:  ||||| : ||::  |     || :| ||||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
                :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
I61902
MHC class I histocompatibility antigen HLA-A alpha chain precursor - human (isolate A*0
C;Species: Homo sapiens (man)
A;Variety: isolate A*0212
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C;Accession: I61902
R;Belich, M.P.; Madrigal, J.A.; Hildebrand, W.H.; Zemmour, J.; Williams, R.C.; Luz, R.;
Nature 357, 326-329, 1992
A;Title: Unusual HLA-B alleles in two tribes of Brazilian Indians.
A;Reference number: I37120; MUID:92269955; PMID:1317015
A;Accession: I61902
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M84378; NID:g187625; PIDN:AAA59604.1; PID:g187626
A;Experimental source: cell line KRC 033; isolate A*0212
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: MHC class I histocompatibility antigen HLA-A alpha chain #status pred
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.5%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 7.9e-35;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |    |    |   | |:|||||  || :| :  |:|:||| ||:     : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|      ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
                :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: August  5, 2003, 13:10:36
Job time : 16 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:05:29 ; Search time 9.5 Seconds
(without alignments)
1366.250 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1513	99.5	348	1	HFE_HUMAN	Q30201 homo sapien
2	1238	81.4	348	1	HFE_DICSU	Q9g142 dicerorhinu
3	1236	81.3	348	1	HFE_CERSI	Q9gkz0 ceratotheri
4	1232	81.1	348	1	HFE_RHIUN	Q9g141 rhinoceros
5	1229	80.9	348	1	HFE_DICBI	Q9g143 diceros bic
6	1156	76.1	360	1	HFE_RAT	O35799 rattus norv
7	1140	75.0	359	1	HFE_MOUSE	P70387 mus musculu
8	523	34.4	361	1	HA1A_RABIT	P01894 oryctolagus
9	523	34.4	361	1	HA1B_RABIT	P06140 oryctolagus
10	517	34.0	365	1	1A01_PANTR	P16209 pan troglod
11	516	33.9	364	1	HA1B_BOVIN	P13753 bos taurus
12	514	33.8	365	1	1A11_HUMAN	P13746 homo sapien
13	512	33.7	370	1	1A03_HUMAN	P04439 homo sapien
14	510	33.6	365	1	1A80_HUMAN	Q09160 homo sapien
15	508	33.4	365	1	1A31_HUMAN	P16189 homo sapien
16	506	33.3	365	1	1A02_HUMAN	P01892 homo sapien
17	506	33.3	365	1	1A30_HUMAN	P16188 homo sapien
18	506	33.3	365	1	1A74_HUMAN	P30459 homo sapien
19	504	33.2	365	1	1A03_PANTR	P13748 pan troglod
20	503	33.1	365	1	1A33_HUMAN	P16190 homo sapien
21	503	33.1	365	1	1A36_HUMAN	P30455 homo sapien
22	503	33.1	365	1	1A68_HUMAN	P01891 homo sapien
23	501.5	33.0	362	1	HA19_CANFA	P18466 canis famil
24	501	33.0	365	1	1A01_HUMAN	P30443 homo sapien
25	500	32.9	273	1	1A69_HUMAN	P10316 homo sapien
26	500	32.9	365	1	1A04_PANTR	P13749 pan troglod
27	500	32.9	365	1	1A24_HUMAN	P05534 homo sapien
28	498	32.8	360	1	HA1A_BOVIN	P13752 bos taurus
29	497	32.7	296	1	ZA2G_RAT	Q63678 rattus norv
30	497	32.7	362	1	1B45_HUMAN	P30485 homo sapien
31	496	32.6	365	1	1A23_HUMAN	P30447 homo sapien
32	494	32.5	338	1	1B20_HUMAN	P30467 homo sapien
33	493	32.4	363	1	1B04_GORGO	P30382 gorilla gor
34	492	32.4	295	1	ZA2G_HUMAN	P25311 homo sapien
35	492	32.4	322	1	HA10_MOUSE	P01898 mus musculu
36	492	32.4	362	1	1B29_HUMAN	P18463 homo sapien
37	492	32.4	371	1	HA12_RAT	P16391 rattus norv
38	491	32.3	365	1	1A34_HUMAN	P30453 homo sapien
39	491	32.3	365	1	1A66_HUMAN	P30457 homo sapien
40	490	32.2	338	1	HLAG_HUMAN	P17693 homo sapien
41	490	32.2	361	1	1B14_HUMAN	P03989 homo sapien
42	490	32.2	362	1	1B18_HUMAN	P10318 homo sapien
43	490	32.2	366	1	1C02_GORGO	P30385 gorilla gor
44	490	32.2	366	1	1C04_GORGO	P30387 gorilla gor
45	489	32.2	359	1	1B01_PANTR	P13750 pan troglod

ALIGNMENTS

RESULT 1
HFE_HUMAN
ID HFE_HUMAN STANDARD; PRT; 348 AA.
AC Q30201; O75929; O75930; O75931; Q96KU5; Q96KU7; Q96KU8; Q9HC64;
AC Q9HC68; Q9HC70; Q9HC83;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hereditary hemochromatosis protein precursor (HLA-H).
GN HFE OR HLAH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=96331279; PubMed=8696333;
RA Feder J.N., Gnirke A., Thomas W., Tsuchihashi Z., Ruddy D.A.,
RA Basava A., Dormishian F., Domingo R., Ellis M.C. Jr., Fullan A.,
RA Hinton L.M., Jones N.L., Kimmel B.E., Kronmal G.S., Lauer P.,
RA Lee V.K., Loeb D.B., Mapa F.A., McClelland E., Meyer N.C.,
RA Mintier G.A., Moeller N., Moore T., Morikang E., Prass C.E.,
RA Quintana L., Starnes S.M., Schatzman R.C., Brunke K.J.,
RA Drayna D.T., Risch N.J., Bacon B.R., Wolff R.K.;
RT "A novel MHC class I-like gene is mutated in patients with hereditary
RT haemochromatosis.";
RL Nat. Genet. 13:399-409(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Albig W., Burmester N., Bode C., Doenecke D., Drabent B.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97294057; PubMed=9149941;
RA Ruddy D.A., Kronmal G.S., Lee V.K., Mintier G.A., Quintana L.,
RA Domingo R. Jr., Meyer N.C., Irrinki A., McClelland E.E., Fullan A.,
RA Mapa F.A., Moore T., Thomas W., Loeb D.B., Harmon C., Tsuchihashi Z.,
RA Wolff R.K., Schatzman R.C., Feder J.N.;
RT "A 1.1-Mb transcript map of the hereditary hemochromatosis locus.";
RL Genome Res. 7:441-456(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Gasparini P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).
RX MEDLINE=99180629; PubMed=10079302;
RA Rhodes D.A., Trowsdale J.;
RT "Alternate splice variants of the hemochromatosis gene Hfe.";
RL Immunogenetics 49:357-359(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 5; 6 AND 7).
RA Oliva R., Sanchez M.;
RT "Identification of different alternative splicing forms of the HFE
RT gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.

RN [7]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 7; 8; 9 AND 10).
RX MEDLINE=20448010; PubMed=11001625;
RA Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.-Y.,
RA David V., Mosser J.;
RT "The HFE gene undergoes alternate splicing processes.";
RL Blood Cells Mol. Dis. 26:155-162(2000).
RN [8]
RP FUNCTION.
RX MEDLINE=98132614; PubMed=9465039;
RA Feder J.N., Penny D.M., Irrinki A., Lee V.K., Lebron J.A., Watson N.,
RA Tsuchihashi Z., Sigal E., Bjorkman P.J., Schatzman R.C.;
RT "The hemochromatosis gene product complexes with the transferrin
RT receptor and lowers its affinity for ligand binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1472-1477(1998).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=98206473; PubMed=9546397;
RA Lebron J.A., Bennett M.J., Vaughn D.E., Chirino A.J., Snow P.M.,
RA Mintier G.A., Feder J.N., Bjorkman P.J.;
RT "Crystal structure of the hemochromatosis protein HFE and
RT characterization of its interaction with transferrin receptor.";
RL Cell 93:111-123(1998).
RN [10]
RP VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=97260408; PubMed=9106528;
RA Carella M., D'Ambrosio L., Totaro A., Grifa A., Valentino M.A.,
RA Piperno A., Girelli D., Roetto A., Franco B., Gasparini P.,
RA Camaschella C.;
RT "Mutation analysis of the HLA-H gene in Italian hemochromatosis
RT patients.";
RL Am. J. Hum. Genet. 60:828-832(1997).
RN [11]
RP VARIANT HH/PCT TYR-282.
RX MEDLINE=97176837; PubMed=9024376;
RA Roberts A.G., Whatley S.D., Morgan R.R., Worwood M., Elder G.H.;
RT "Increased frequency of the haemochromatosis Cys282Tyr mutation in
RT sporadic porphyria cutanea tarda.";
RL Lancet 349:321-323(1997).
RN [12]
RP VARIANT HH/PCT ASP-63.
RX MEDLINE=98085904; PubMed=9425935;
RA Sampietro M., Piperno A., Lupica L., Arosio C., Vergani A.,
RA Corbetta N., Malosio I., Mattioli M., Fracanzani A.L.,
RA Cappellini M.D., Fiorelli G., Fargion S.;
RT "High prevalence of the His63Asp HFE mutation in Italian patients with
RT porphyria cutanea tarda.";
RL Hepatology 27:181-184(1998).
RN [13]
RP VARIANTS HH/PCT ASP-63 AND TYR-282.
RX MEDLINE=98281650; PubMed=9620340;
RA Bonkovsky H.L., Poh-Fitzpatrick M., Pimstone N., Obando J.,
RA Di Bisceglie A., Tattrie C., Tortorelli K., LeClair P., Mercurio M.G.,
RA Lambrecht R.W.;
RT "Porphyria cutanea tarda, hepatitis C, and HFE gene mutations in North
RT America.";
RL Hepatology 27:1661-1669(1998).
RN [14]
RP VARIANTS HH ASP-63; CYS-65 AND TYR-282.
RX MEDLINE=99211934; PubMed=10194428;
RA Mura C., Raguenes O., Ferec C.;
RT "HFE mutations analysis in 711 hemochromatosis probands: evidence for
RT S65C implication in mild form of hemochromatosis.";
RL Blood 93:2502-2505(1999).
RN [15]
RP VARIANTS HH CYS-65; ARG-93 AND THR-105.
RX MEDLINE=20042794; PubMed=10575540;
RA Barton J.C., Sawada-Hirai R., Rothenberg B.E., Acton R.T.;
RT "Two novel missense mutations of the HFE gene (I105T and G93R) and
RT identification of the S65C mutation in Alabama hemochromatosis
RT probands.";
RL Blood Cells Mol. Dis. 25:147-155(1999).
RN [16]
RP VARIANTS VP ASP-63 AND HIS-127, VARIANT HH MET-330, AND VARIANTS
RP MET-53 AND MET-59.
RX MEDLINE=99330560; PubMed=10401000;
RA de Villiers J.N.P., Hillermann R., Loubser L., Kotze M.J.;
RT "Spectrum of mutations in the HFE gene implicated in haemochromatosis
RT and porphyria.";
RL Hum. Mol. Genet. 8:1517-1522(1999).
RN [17]
RP VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=99140260; PubMed=10094552;
RA Merryweather-Clarke A.T., Simonsen H., Shearman J.D., Pointon J.J.,
RA Norgaard-Pedersen B., Robson K.J.H.;
RT "A retrospective anonymous pilot study in screening newborns for HFE
RT mutations in Scandinavian populations.";
RL Hum. Mutat. 13:154-159(1999).
RN [18]
RP VARIANT HH CYS-65.
RA Fagan E., Payne S.J.;
RT "A novel missense mutation S65C in the HFE gene with a possible role
RT in hereditary haemochromatosis.";
RL Hum. Mutat. 13:507-508(1999).
RN [19]
RP VARIANT LYS-277.
RX MEDLINE=20081073; PubMed=10612845;
RA Bradbury R., Fagan E., Payne S.J.;
RT "Two novel polymorphisms (E277K and V212V) in the haemochromatosis
RT gene HFE.";
RL Hum. Mutat. 15:120-120(2000).
CC -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC affinity for iron-loaded transferrin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q30201-1; Sequence=Displayed;
CC Name=2; Synonyms=delE2;
CC IsoId=Q30201-2; Sequence=VSP_003218;
CC Name=3; Synonyms=del14E4;
CC IsoId=Q30201-3; Sequence=VSP_003225;
CC Name=4; Synonyms=delE214E4;
CC IsoId=Q30201-4; Sequence=VSP_003218, VSP_003225;
CC Name=5;
CC IsoId=Q30201-5; Sequence=VSP_003219;
CC Name=6;
CC IsoId=Q30201-6; Sequence=VSP_003220;
CC Name=7; Synonyms=delE3;
CC IsoId=Q30201-7; Sequence=VSP_003221;
CC Name=8; Synonyms=1043-2283del,intron6ins;
CC IsoId=Q30201-8; Sequence=VSP_003226, VSP_003227;
CC Name=9; Synonyms=delE3-7;
CC IsoId=Q30201-9; Sequence=VSP_003223, VSP_003224;
CC Name=10; Synonyms=562-878del;
CC IsoId=Q30201-10; Sequence=VSP_003222;
CC -!- TISSUE SPECIFICITY: IN ALL TISSUES TESTED EXCEPT BRAIN.
CC -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF HEREDITARY HEMOCHROMATOSIS
CC (HH). HH IS AN AUTOSOMAL RECESSIVE INBORN DISORDER OF IRON
CC METABOLISM, FREQUENT AMONG CAUCASIANS. HH IS CHARACTERIZED BY
CC ABNORMAL INTESTINAL IRON ABSORPTION AND PROGRESSIVE INCREASE OF
CC TOTAL BODY IRON, WHICH RESULTS IN MIDLIFE IN CLINICAL
CC COMPLICATIONS INCLUDING CIRRHOSIS, CARDIOPATHY, DIABETES,
CC ENDOCRINE DYSFUNCTIONS, ARTHROPATHY, AND SUSCEPTIBILITY TO LIVER
CC CANCER. SINCE THE DISEASE COMPLICATIONS CAN BE EFFECTIVELY
CC PREVENTED BY REGULAR PHLEBOTOMIES, EARLY DIAGNOSIS IS MOST
CC IMPORTANT TO PROVIDE A NORMAL LIFE EXPECTANCY TO THE AFFECTED
CC SUBJECTS.
CC -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF PORPHYRIA CUTANEA TARDA
CC (PCT), A DISORDER CHARACTERIZED BY LIGHT-SENSITIVE DERMATITIS AND
CC PRESENCE OF LARGE AMOUNTS OF UROPORPHYRIN IN URINE. IRON OVERLOAD
CC IS OFTEN PRESENT IN ASSOCIATION WITH VARYING DEGREES OF LIVER
CC DAMAGE. PCT IS THE MOST COMMON FORM OF PORPHYRIA WORLDWIDE. IT
CC OCCURS IN TWO FORMS: THE SPORADIC TYPE (PCT TYPE I) AND THE
CC FAMILIAL TYPE (PCT TYPE II), WHICH IS INHERITED IN AN AUTOSOMAL

```
  Query Match             99.5%;  Score 1513;  DB 1;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.3e-119;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298
```

RESULT 2
HFE_DICSU
ID HFE_DICSU STANDARD; PRT; 348 AA.
AC Q9GL42;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hereditary hemochromatosis protein precursor.
GN HFE.
OS Dicerorhinus sumatrensis (Sumatran rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.
OX NCBI_TaxID=89632;
RN [1]
RP SEQUENCE FROM N.A.
RA West C.J., Worley M., Beutler E.;
RT "Rhinoceros HFE polymorphisms.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC affinity for iron-loaded transferrin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC ---

CC ---
DR EMBL; AY007543; AAG23703.1; -.
DR HSSP; Q30201; 1A6Z.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 348 HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT DOMAIN 23 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 205 EXTRACELLULAR ALPHA-2.
FT DOMAIN 206 297 EXTRACELLULAR ALPHA-3.
FT DOMAIN 298 306 CONNECTING PEPTIDE.
FT TRANSMEM 307 330 POTENTIAL.
FT DOMAIN 331 348 CYTOPLASMIC TAIL.
FT DISULFID 124 187 BY SIMILARITY.
FT DISULFID 225 282 BY SIMILARITY.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 348 AA; 39740 MW; 518BFD357AB83B90 CRC64;

```
  Query Match             81.4%;  Score 1238;  DB 1;  Length 348;
  Best Local Similarity   81.3%;  Pred. No. 1.5e-96;
  Matches  222;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||||:| || ||||||||||:||| |: |||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              |||||||||||||| |||||||:||||||||||||||||:|||||| |:||||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |: |||||||||||||||||||||||||   ||
Db        266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298
```

RESULT 3
HFE_CERSI
ID HFE_CERSI STANDARD; PRT; 348 AA.
AC Q9GKZ0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hereditary hemochromatosis protein precursor.
GN HFE.
OS Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX NCBI_TaxID=9807;
RN [1]
RP SEQUENCE FROM N.A.
RA West C.J., Worley M., Beutler E.;
RT "Rhinoceros HFE polymorphisms.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC affinity for iron-loaded transferrin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC ---

CC ---
DR EMBL; AY007541; AAG23701.1; -.
DR HSSP; Q30201; 1A6Z.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.

```
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39822 MW;  2523016ECE9FBE91 CRC64;

  Query Match             81.3%;  Score 1236;  DB 1;  Length 348;
  Best Local Similarity   81.7%;  Pred. No. 2.1e-96;
  Matches  223;  Conservative   18;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |   ||      ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              |||||||||||| |||||||:||||||||||||||||:|||||| |:||||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:||||||:|||:|||||||:||||::||| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPMDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
               |  ||||||||||||||||||||||||   ||
Db        266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 4
HFE_RHIUN
ID   HFE_RHIUN      STANDARD;      PRT;   348 AA.
AC   Q9GL41;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Rhinoceros unicornis (Greater Indian rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
OX   NCBI_TaxID=9809;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007544; AAG23704.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39743 MW;  F2723D57A327A6B4 CRC64;

  Query Match             81.1%;  Score 1232;  DB 1;  Length 348;
  Best Local Similarity   81.0%;  Pred. No. 4.6e-96;
  Matches  221;  Conservative   20;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |   ||      ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              |||||||||||| |||||||:||||||||||||||||:|||||| |:| ||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWMYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
               |: ||||||||||||||||||||||   ||
Db        266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 5
HFE_DICBI
ID   HFE_DICBI      STANDARD;      PRT;   348 AA.
AC   Q9GL43; Q9GK81;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
```

```
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007542; AAG23702.1; -.
DR   EMBL; AF301592; AAG39940.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39818 MW;  4D95E7B01E48FB90 CRC64;

  Query Match             80.9%;  Score 1229;  DB 1;  Length 348;
  Best Local Similarity   81.0%;  Pred. No. 8.3e-96;
  Matches  221;  Conservative   20;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              ||:|||||||||| |||||||:||||||||||||||||:|||||| |:||||||||||
Db         86 QLTQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||||| ||| ||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |  |||||||||||||||||||||||   ||
Db        266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298
```

```
RESULT 6
HFE_RAT
ID   HFE_RAT        STANDARD;      PRT;   360 AA.
AC   O35799; O35175;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor (RT1-CAFE).
GN   HFE.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Banasch M.W., Schaefer H., Schmidt W.E.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 42-303 FROM N.A.
RC   TISSUE=Small intestine;
RA   Sawada-Hirai R., Rothenberg B.E.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS TO TRANSFERRIN RECEPTOR (TFR) AND REDUCES ITS
CC       AFFINITY FOR IRON-LOADED TRANSFERRIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001517; CAA04799.1; -.
DR   EMBL; AF008587; AAB86597.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    360       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       26    127       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      128    218       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      219    310       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      311    319       CONNECTING PEPTIDE.
FT   TRANSMEM    320    340       POTENTIAL.
FT   DOMAIN      341    360       CYTOPLASMIC TAIL.
FT   DISULFID    137    200       BY SIMILARITY.
FT   DISULFID    238    295       BY SIMILARITY.
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    198    198       R -> K (IN REF. 2).
SQ   SEQUENCE   360 AA;  40988 MW;  CC819834EE3240B3 CRC64;

  Query Match             76.1%;  Score 1156;  DB 1;  Length 360;
  Best Local Similarity   73.6%;  Pred. No. 1.1e-89;
  Matches  206;  Conservative   29;  Mismatches   37;  Indels    8;  Gaps    1;
```

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQ 64
              |||| |||||||: |||| ||||||||||||| |: |||| ||| ||:  : |||:|||
Db         32 SHSLRYLFMGASKPDLGLPFFEALGYVDDQLFVSYNHESRRAEPRAPWILGQTSSQLWLQ 91

Qy         65 LSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWKY 116
              |||||||||:|| ||||||| |:|||        ||| ||||||||: ||||| |:|||
Db         92 LSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLRVVPESHILQVILGCEVHEDNSTSGFWKY 151

Qy        117 GYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELG 176
              |||||||||||| ||:| ||||||| ||:||| |:|||||:| ||:|||| ||:|:|||
Db        152 GYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHRIRARQSRDYLQRDCPQQLKQVLELQ 211

Qy        177 RGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              |||| |||| ||||| |  |: |:|||:|||::|||||:|||| ||:|||:  |::|||
Db        212 RGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQNITMRWLKDSQPLDAKDVNPENVLP 271

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||:|||| |||| |::|||||||||||||    ||
Db        272 NGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPLTATWE 311


RESULT 7
HFE_MOUSE
ID   HFE_MOUSE      STANDARD;      PRT;   359 AA.
AC   P70387;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor.
GN   HFE OR MR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98060831; PubMed=9396865;
RA   Riegert P., Gilfillan S., Nanda I., Schmid M., Bahram S.;
RT   "The mouse HFE gene.";
RL   Immunogenetics 47:174-177(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Lung;
RA   Hashimoto K.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 37-211 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=97148566; PubMed=9020055;
RA   Hashimoto K., Hirai M., Kurosawa Y.;
RT   "Identification of a mouse homolog for the human hereditary
RT   haemochromatosis candidate gene.";
RL   Biochem. Biophys. Res. Commun. 230:35-39(1997).
RN   [4]
RP   SEQUENCE OF 79-359 FROM N.A.
RC   STRAIN=129;
RA   Albig W., Drabent B., Doenecke D.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF007558; AAC03447.1; -.
DR   EMBL; U66849; AAB07525.1; -.
DR   EMBL; Y12650; CAA73197.1; -.
DR   EMBL; U80604; AAB51504.1; -.
DR   PIR; JC5382; JC5382.
DR   HSSP; Q30201; 1A6Z.
DR   MGD; MGI:109191; Hfe.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    359       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       25    126       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      127    217       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      218    309       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      310    318       CONNECTING PEPTIDE.
FT   TRANSMEM    319    339       POTENTIAL.
FT   DOMAIN      340    359       CYTOPLASMIC TAIL.
FT   DISULFID    136    199       BY SIMILARITY.
FT   DISULFID    237    294       BY SIMILARITY.
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   359 AA;  40548 MW;  4BDE6C27F9FF20B4 CRC64;

  Query Match             75.0%;  Score 1140;  DB 1;  Length 359;
  Best Local Similarity   72.2%;  Pred. No. 2.5e-88;
  Matches  203;  Conservative   30;  Mismatches   40;  Indels    8;  Gaps    1;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |: |||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              |||||||||:|| ||||||| |:||||        ||| |||:||||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:||||||||||||    ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310


RESULT 8
HA1A_RABIT
ID   HA1A_RABIT     STANDARD;      PRT;   361 AA.
AC   P01894;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 11/11 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84290724; PubMed=6432910;
RA   Tykocinski M.L., Marche P.N., Max E.E., Kindt T.J.;
RT   "Rabbit class I MHC genes: cDNA clones define full-length transcripts
RT   of an expressed gene and a putative pseudogene.";
RL   J. Immunol. 133:2261-2269(1984).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02441; AAA98729.1; -.
DR   PIR; A02193; HLRB.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN 11/11.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    329
FT   DOMAIN      330    361       CYTOPLASMIC.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40447 MW;  580B673323C1AE35 CRC64;

  Query Match            34.4%;  Score 523;  DB 1;  Length 361;
  Best Local Similarity  40.1%;  Pred. No. 1.2e-36;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |      | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::    |  | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   ||||  |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

```
RESULT 9
HA1B_RABIT
ID   HA1B_RABIT     STANDARD;      PRT;   361 AA.
AC   P06140;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 19-1 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85103547; PubMed=3917974;
RA   Marche P.N., Tykocinski M.L., Max E.E., Kindt T.J.;
RT   "Structure of a functional rabbit class I MHC gene: similarity to
RT   human class I genes.";
RL   Immunogenetics 21:71-82(1985).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02819; AAA98730.1; -.
DR   PIR; I46858; I46858.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN 19-1.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    329
FT   DOMAIN      330    361       CYTOPLASMIC.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40455 MW;  C06FBD8B87ED0546 CRC64;

  Query Match            34.4%;  Score 523;  DB 1;  Length 361;
  Best Local Similarity  40.1%;  Pred. No. 1.2e-36;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |      | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144
```

```
Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::    |   | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
                :  ||   ||||  |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 10
1A01_PANTR
ID   1A01_PANTR     STANDARD;      PRT;   365 AA.
AC   P16209;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   CHLA class I histocompatibility antigen, A-2 alpha chain precursor.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90201944; PubMed=1690682;
RA   Lawlor D.A., Warren E., Ward F.E., Parham P.;
RT   "Comparison of class I MHC alleles in humans and apes.";
RL   Immunol. Rev. 113:147-185(1990).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30678; AAA87970.1; -.
DR   PIR; I36961; I36961.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-2 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   365 AA;  40848 MW;  FC452786BD038D3E CRC64;
```

```
  Query Match            34.0%;  Score 517;  DB 1;  Length 365;
  Best Local Similarity  39.7%;  Pred. No. 3.9e-36;
  Matches  110;  Conservative   45;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |    |     |    | |:|||||  || :| :  |:|:||| ||:       : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
                : ::| |          ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
                :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 11
HA1B_BOVIN
ID   HA1B_BOVIN     STANDARD;      PRT;   364 AA.
AC   P13753;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   BOLA class I histocompatibility antigen, alpha chain BL3-7 precursor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258075; PubMed=3133413;
RA   Ennis P.D., Jackson A.P., Parham P.;
RT   "Molecular cloning of bovine class I MHC cDNA.";
RL   J. Immunol. 141:642-651(1988).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21043; AAA30641.1; -.
DR   HSSP; P16391; 1ED3.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28    364       BOLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
```

```
FT                                ALPHA CHAIN BL3-7.
FT   DOMAIN       28    117       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      118    209       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      210    301       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      302    310       CONNECTING PEPTIDE.
FT   TRANSMEM    311    331
FT   DOMAIN      332    364       CYTOPLASMIC.
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    128    191       BY SIMILARITY.
FT   DISULFID    230    286       BY SIMILARITY.
SQ   SEQUENCE   364 AA;  41513 MW;  622056CF7DCF7873 CRC64;

  Query Match             33.9%;  Score 516;  DB 1;  Length 364;
  Best Local Similarity   38.9%;  Pred. No. 4.8e-36;
  Matches  109;  Conservative   50;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISS 59
              |  |||| | : | |    ||    | |:|||||  || :| :  : | ||| ||:
Db         26 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 84

Qy         60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
              : | : ::  |    :| || |:   :| |:  || :| : ||::  |     |:|::|
Db         85 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 144

Qy        118 YDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
              |||:|::    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db        145 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 203

Qy        178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              | :  ||   |||| ::    |||| || :|| |:: | :: :    :: |  :  |
Db        204 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 262

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              :||||:| |  | || ||||||||:|:| || :|| : ||
Db        263 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 302


RESULT 12
1A11_HUMAN
ID   1A11_HUMAN     STANDARD;      PRT;   365 AA.
AC   P13746;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HLA class I histocompatibility antigen, A-11 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*1101/A*1102).
RX   MEDLINE=89030641; PubMed=2460344;
RA   Mayer W.E., Jonker M., Klein D., Ivanyi P., van Seventer G.,
RA   Klein J.;
RT   "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for
RT   trans-species mode of evolution.";
RL   EMBO J. 7:2765-2774(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (A*1101/A*1102).
RX   MEDLINE=94287401; PubMed=8016845;
RA   Lin L., Tokunaga K., Ishikawa Y., Bannai M., Kashiwase K.,
RA   Kuwata S., Akaza T., Tadokoro K., Shibata Y., Juji T.;
RT   "Sequence analysis of serological HLA-A11 split antigens, A11.1 and
RT   A11.2.";
RL   Tissue Antigens 43:78-82(1994).
RN   [3]
RP   SEQUENCE OF 26-365 FROM N.A. (A*1101).
RX   MEDLINE=87192928; PubMed=2437024;
RA   Cowan E.P., Jelachich M.L., Biddison W.E., Coligan J.E.;
RT   "DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows
RT   identification of residues involved in epitopes recognized by
RT   antibodies and T cells.";
RL   Immunogenetics 25:241-250(1987).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-11 ARE KNOWN: A*1101 (A-
CC       11E) AND A*1102 (A-11K). THE SEQUENCE SHOWN IS THAT OF A*1101.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13111; CAA31503.1; -.
DR   EMBL; X13112; CAA31504.1; -.
DR   EMBL; D16841; BAA04117.1; -.
DR   EMBL; D16842; BAA04118.1; -.
DR   EMBL; M16010; AAA65449.1; -.
DR   EMBL; M16007; AAA65449.1; JOINED.
DR   EMBL; M16008; AAA65449.1; JOINED.
DR   EMBL; M16009; AAA65449.1; JOINED.
DR   PIR; I83063; I83063.
DR   PIR; S03536; A47636.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-11 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   VARIANT      43     43       E -> K (IN ALLELE A*1102).
FT                                /FTId=VAR_004353.
SQ   SEQUENCE   365 AA;  40937 MW;  FE449CE2D4BF6CC5 CRC64;

  Query Match             33.8%;  Score 514;  DB 1;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 7e-36;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
```

```
           :|::    |   | ||:  |  || :||     | | |||||  |  . |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 13
1A03_HUMAN
ID   1A03_HUMAN     STANDARD;      PRT;   370 AA.
AC   P04439;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HLA class I histocompatibility antigen, A-3 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*0301).
RX   MEDLINE=84207948; PubMed=6609814;
RA   Strachan T., Sodoyer R., Damotte M., Jordan B.R.;
RT   "Complete nucleotide sequence of a functional class I HLA gene,
RT   HLA-A3: implications for the evolution of HLA genes.";
RL   EMBO J. 3:887-894(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (A*0302).
RX   MEDLINE=85290871; PubMed=2993417;
RA   Cowan E.P., Jordan B.E., Coligan J.E.;
RT   "Molecular cloning and DNA sequence analysis of genes encoding
RT   cytotoxic T lymphocyte-defined HLA-A3 subtypes: the E1 subtype.";
RL   J. Immunol. 135:2835-2841(1985).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-3 ARE KNOWN: A*0301 (A-
CC       3.1) AND A*0302. THE SEQUENCE SHOWN IS THAT OF A*0301.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00492; CAA25162.1; ALT_TERM.
DR   PIR; A02192; HLHUA3.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     29
FT   CHAIN         30    370       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-3 ALPHA CHAIN.
FT   DOMAIN        30    119       EXTRACELLULAR ALPHA-1.
FT   DOMAIN        120   211       EXTRACELLULAR ALPHA-2.
FT   DOMAIN        212   303       EXTRACELLULAR ALPHA-3.
FT   DOMAIN        304   313       CONNECTING PEPTIDE.
FT   TRANSMEM      314   337
FT   DOMAIN        338   370       CYTOPLASMIC TAIL.
FT   CARBOHYD      115   115       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID      130   193       BY SIMILARITY.
FT   DISULFID      232   288       BY SIMILARITY.
FT   VARIANT       181   181       E -> V (IN ALLELE A*0302).
FT                                 /FTId=VAR_004351.
FT   VARIANT       185   185       L -> Q (IN ALLELE A*0302).
FT                                 /FTId=VAR_004352.
SQ   SEQUENCE   370 AA;  41368 MW;  ABB1FA77460318A2 CRC64;

  Query Match             33.7%;  Score 512;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 1.1e-35;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels   10;  Gaps    8;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
                | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db         90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              :|::    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db        150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy        180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db        208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||:| |  : || |||||||| |:| || :|| : ||
Db        267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 14
1A80_HUMAN
ID   1A80_HUMAN     STANDARD;      PRT;   365 AA.
AC   Q09160;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, AW-80(A-1) alpha chain
DE   precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94245293; PubMed=8188325;
RA   Balas A., Garcia-Sanchez F., Gomez-Reino F., Vicario J.L.;
RT   "Characterization of a new and highly distinguishable HLA-A allele in
RT   a Spanish family.";
RL   Immunogenetics 39:452-452(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Domena J.D.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE ONLY ALLELE OF AW-80 KNOWN IS A*8001 WHICH IS
```

```
CC       SHOWN HERE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U03754; AAC04322.1; -.
DR   EMBL; L18898; AAA17012.1; -.
DR   PIR; I59638; I38439.
DR   HSSP; Q95352; 1HHK.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 AW-80(A-1) ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   365 AA;  40791 MW;  CE1BC1CD60CA8FA8 CRC64;

  Query Match           33.6%;  Score 510;  DB 1;  Length 365;
  Best Local Similarity  38.3%;  Pred. No. 1.5e-35;
  Matches  106;  Conservative   53;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::::|        :  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||  : || | ||||| :|   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| ||:||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|:|||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 15
1A31_HUMAN
ID   1A31_HUMAN     STANDARD;      PRT;   365 AA.
AC   P16189; O98137; Q9TQ24;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HLA class I histocompatibility antigen, A-31 alpha chain precursor
DE   (Aw-19).
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*3101).
RX   MEDLINE=90038496; PubMed=2478623;
RA   Kato K., Trapani J.A., Allopenna J., Dupont B., Yang S.Y.;
RT   "Molecular analysis of the serologically defined HLA-Aw19 antigens. A
RT   genetically distinct family of HLA-A antigens comprising A29, A31,
RT   A32, and Aw33, but probably not A30.";
RL   J. Immunol. 143:3371-3378(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (A*3101).
RX   MEDLINE=92269955; PubMed=1317015;
RA   Belich M.P., Madrigal J.A., Hildebrand W.H., Zemmour J.,
RA   Williams R.C., Luz R., Petzl-Erler M.L., Parham P.;
RT   "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL   Nature 357:326-329(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (A*31012).
RX   MEDLINE=96387675; PubMed=8795145;
RA   Arnett K.L., Adams E.J., Parham P.;
RT   "On the sequence of A*3101.";
RL   Tissue Antigens 47:428-430(1996).
RN   [4]
RP   SEQUENCE OF 9-365 FROM N.A. (A*3101).
RX   MEDLINE=92269956; PubMed=1589035;
RA   Watkins D.I., McAdam S.N., Liu X., Stang C.R., Milford E.L.,
RA   Levine C.G., Garber T.L., Dogon A.L., Lord C.I., Ghim S.H.,
RA   Troup G.M., Hughes A.L., Letvin N.L.;
RT   "New recombinant HLA-B alleles in a tribe of South American
RT   Amerindians indicate rapid evolution of MHC class I loci.";
RL   Nature 357:329-333(1992).
RN   [5]
RP   SEQUENCE FROM N.A. (A*3104).
RA   Bettinotti M.P., Dhillon G., Hackett J., Simonis T.B., Marincola F.M.;
RT   "A New HLA-A*31 allele.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 26-206 FROM N.A. (A*3104).
RA   Mitsuishi Y.;
RT   "New HLA-A31 allele identified in African American population.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-31 ARE KNOWN: A*3101 AND
CC       A*3104. THE SEQUENCE SHOWN IS THAT OF A*3101.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30578; AAA59613.1; -.
DR   EMBL; M84375; AAA59599.1; -.
DR   EMBL; L78918; AAB05976.1; -.
DR   EMBL; AF148863; AAD39981.1; -.
DR   EMBL; AF105028; AAC79721.1; -.
DR   EMBL; AF105027; AAC79721.1; JOINED.
DR   PIR; I72170; I72170.
DR   HSSP; O19673; 1HSB.
```

```
DR   MIM; 142800; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-31 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   VARIANT     121    121       M -> I (IN ALLELE A*3104).
FT                                 /FTId=VAR_010373.
FT   VARIANT     138    138       Q -> R (IN ALLELE A*3104).
FT                                 /FTId=VAR_010374.
SQ   SEQUENCE   365 AA;  41004 MW;  4E760C821A3C553B CRC64;

  Query Match             33.4%;  Score 508;  DB 1;  Length 365;
  Best Local Similarity   39.0%;  Pred. No. 2.2e-35;
  Matches  108;  Conservative   51;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: |     |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFTTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|    :  ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAHSQIDRVDLGTLRGYYNQSEAGSHTIQMMYGCDVGSDGRFLRGYQQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  |: :||  :: | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITQRKWEAARV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| | :: || |:|||||| |:| || :|| : ||
Db        263 GTFQKWASVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: August  5, 2003, 13:08:48
Job time : 10.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:05:29 ; Search time 32 Seconds
(without alignments)
2225.704 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL...:.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1140	75.0	358	11	Q8C2A6	Q8c2a6 mus musculu
2	1140	75.0	359	11	Q9D754	Q9d754 mus musculu
3	802	52.8	272	11	Q9R105	Q9r105 rattus norv
4	592	38.9	116	4	Q9HC69	Q9hc69 homo sapien
5	547.5	36.0	359	7	Q8HX81	Q8hx81 ornithorhyn
6	543.5	35.8	340	7	Q9BD50	Q9bd50 pongo pygma
7	542.5	35.7	334	7	Q9TQK3	Q9tqk3 homo sapien
8	542.5	35.7	341	4	Q9NPL2	Q9npl2 homo sapien
9	542.5	35.7	341	7	Q95460	Q95460 homo sapien
10	542.5	35.7	341	7	Q9BCU3	Q9bcu3 pan troglod
11	540.5	35.6	354	7	Q95HB3	Q95hb3 anas platyr
12	539.5	35.5	341	7	Q9BCU4	Q9bcu4 pan troglod
13	530	34.9	105	4	Q9HC71	Q9hc71 homo sapien
14	521	34.3	356	7	Q8HX66	Q8hx66 sus scrofa
15	520	34.2	332	7	Q30990	Q30990 pan troglod
16	520	34.2	365	7	Q9TPL7	Q9tpl7 pan troglod
17	518	34.1	365	7	Q9TQP6	Q9tqp6 homo sapien
18	518	34.1	371	7	Q9TQP7	Q9tqp7 homo sapien
19	517	34.0	352	7	Q8MHT1	Q8mht1 sus scrofa
20	517	34.0	364	7	O19243	O19243 sus scrofa
21	517	34.0	365	7	Q9MYG4	Q9myg4 homo sapien
22	515	33.9	365	7	Q29747	Q29747 homo sapien
23	514	33.8	273	7	Q95IG6	Q95ig6 homo sapien
24	514	33.8	352	7	Q8SPA9	Q8spa9 sus scrofa
25	514	33.8	361	7	Q8HX63	Q8hx63 sus scrofa
26	514	33.8	364	7	Q8HX61	Q8hx61 sus scrofa
27	514	33.8	365	7	Q9MYI5	Q9myi5 homo sapien
28	513	33.8	330	7	O19356	O19356 macaca mula
29	513	33.8	331	7	O02944	O02944 macaca mula
30	513	33.8	333	7	O98030	O98030 papio anubi
31	513	33.8	333	7	O98031	O98031 papio anubi
32	512	33.7	129	4	Q9UK37	Q9uk37 homo sapien
33	512	33.7	330	7	O02947	O02947 macaca mula
34	512	33.7	330	7	O02946	O02946 macaca mula
35	512	33.7	365	7	O19756	O19756 homo sapien
36	511	33.6	331	7	O02945	O02945 macaca mula
37	511	33.6	357	7	Q30886	Q30886 pan paniscu
38	511	33.6	363	7	Q9MXI5	Q9mxi5 pan troglod
39	511	33.6	363	7	Q9MWK4	Q9mwk4 gorilla gor
40	511	33.6	365	7	Q9MXI6	Q9mxi6 pan troglod
41	511	33.6	365	7	Q9MXM7	Q9mxm7 pan troglod
42	510	33.6	360	7	Q9GJ24	Q9gj24 homo sapien
43	510	33.6	364	7	Q8SPA4	Q8spa4 sus scrofa
44	510	33.6	365	7	Q30900	Q30900 pan paniscu
45	510	33.6	365	7	Q9BCN0	Q9bcn0 homo sapien

ALIGNMENTS

RESULT 1
Q8C2A6
ID Q8C2A6 PRELIMINARY; PRT; 358 AA.
AC Q8C2A6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hemochromatosis.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK088986; BAC40688.1; -.
SQ SEQUENCE 358 AA; 40421 MW; EE88FB6E5AAC844D CRC64;

Query Match 75.0%; Score 1140; DB 11; Length 358;
Best Local Similarity 72.2%; Pred. No. 2.5e-99;
Matches 203; Conservative 30; Mismatches 40; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |: |||| ||| ||:  : |||:||
Db         29 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 88

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              |||||||||:|| ||||||| |:|||         ||| |||:|||| |||| |:|:
Db         89 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 148

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        149 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 208
```

```
Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::||||||:|||| ||:|||:  |: ||
Db        209 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 268

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:|||||||||||||   ||
Db        269 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 309


RESULT 2
Q9D754
ID   Q9D754      PRELIMINARY;      PRT;   359 AA.
AC   Q9D754;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Adult male tongue cDNA, RIKEN full-length enriched library,
DE   clone:2310032M04, full insert sequence.
GN   HFE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Tongue;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AK009581; BAB26373.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   MGD; MGI:109191; Hfe.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   359 AA;  40534 MW;  586657B7F9FF20B4 CRC64;

  Query Match           75.0%;  Score 1140;  DB 11;  Length 359;
  Best Local Similarity   72.2%;  Pred. No. 2.5e-99;
  Matches  203;  Conservative   30;  Mismatches   40;  Indels    8;  Gaps    1;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |: |||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              ||||||||:|| ||||||| |:||||         ||| |||:||||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::||||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:|||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310


RESULT 3
Q9R105
ID   Q9R105      PRELIMINARY;      PRT;   272 AA.
AC   Q9R105;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemochromatosis gene product HFE splice variant delE2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Testis;
RA   Liew Y.-F., Shaw N.-S.;
RT   "Alternative splice variant of the hemochromatosis gene HFE in iron
RT   overloaded rats.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF176534; AAD49965.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   272 AA;  30757 MW;  1D91063CCBEF5502 CRC64;

  Query Match           52.8%;  Score 802;  DB 11;  Length 272;
  Best Local Similarity   75.1%;  Pred. No. 1.4e-67;
  Matches  139;  Conservative   22;  Mismatches   24;  Indels    0;  Gaps    0;

Qy         92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 151
              ||| |||||||: ||||| |:|||||||||||||| ||:| ||||||| ||:||| |:
Db         39 ESHILQVILGCEVHEDNSTSGFWKYGYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHR 98

Qy        152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQ 211
              |||||:| ||:|||| ||:|:||| |||| |||| ||||| |  |: |:|||:|||::||
Db         99 IRARQSRDYLQRDCPQQLKQVLELQRGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQ 158

Qy        212 NITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPL 271
```

```
                 ||||:|||| ||:|||:  |::||||||||||||:|||| |||| |::||||||||||||
Db        159 NITMRWLKDSQPLDAKDVNPENVLPNGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPL 218

Qy        272 IVIWE 276
                 ||
Db        219 TATWE 223
```

RESULT 4
Q9HC69
ID Q9HC69 PRELIMINARY; PRT; 116 AA.
AC Q9HC69;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hemochromatosis splice variant 861-2305del (Fragment).
GN HFE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20448010; PubMed=11001625;
RA Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA Mosser J.;
RT "The HFE gene undergoes alternate splicing processes.";
RL Blood Cells Mol. Dis. 26:155-162(2000).
DR EMBL; AF144241; AAG29576.1; -.
DR HSSP; Q30201; 1A6Z.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
FT NON_TER 1 1
SQ SEQUENCE 116 AA; 13541 MW; AC0333B096A3F47B CRC64;

Query Match 38.9%; Score 592; DB 4; Length 116;
Best Local Similarity 99.1%; Pred. No. 3.4e-48;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         89 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 60

Qy        149 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVT 195
              |||||||||||||||||||||||||||||||||||||||||||||||:
Db         61 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVS 107
```

RESULT 5
Q8HX81
ID Q8HX81 PRELIMINARY; PRT; 359 AA.
AC Q8HX81;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I antigen.
OS Ornithorhynchus anatinus (Duckbill platypus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX NCBI_TaxID=9258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22226695; PubMed=12242589;
RA Miska K.B., Harrison G.A., Hellman L., Miller R.D.;
RT "The major histocompatibility complex in monotremes: an analysis of
RT the evolution of Mhc class I genes across all three mammalian
RT subclasses.";
RL Immunogenetics 54:381-393(2002).
DR EMBL; AY112715; AAM54213.1; -.
SQ SEQUENCE 359 AA; 39847 MW; 3FC701E44FD53B66 CRC64;

Query Match 36.0%; Score 547.5; DB 7; Length 359;
Best Local Similarity 41.7%; Pred. No. 2.2e-43;
Matches 116; Conservative 49; Mismatches 106; Indels 7; Gaps 6;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFV-FYDDESR-RVEPRTPWVSSRISSQM 61
              ||||| | : | |    |:| | |:||||:  || |  |  | | ||||||:      :
Db         25 RSHSLRYFYTGVSRPGSGVSEFTAVGYVDEVQFVRFEGDGGRGRAEPRTPWIRDNEGQEY 84

Qy         62 WLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNST-EGYWKYGYD 119
              | : :| ||    :|  :  | :  :| ||| ||| | : ||:::||::  :|: :||||
Db         85 WDRNTQILKDTAQVFRTNLQTALGYYNQSKEGSHTFQRMYGCQLREDDTPGQGFTQYGYD 144

Qy        120 GQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              |:|::    || | ||:  |  || :||  :  | :::|||| :|   |:: :  |
Db        145 GRDYIALDKATLSWTAADAGALNTKRKWEADRTIAERDKAYLEDECIEWLRKYVGFGGPS 204

Qy        180 LDQQVPPLVKVTHHV--TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPN 237
              | :  || |:|| |         :|||||| :|| :|  || :: : | ::| |      |:
Db        205 LTRADPPEVEVTRHTGPDGEDVSLRCRALGFYPADIKFKWEREGKDM-SQEMEFVGTRPS 263

Qy        238 GDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
              ||| :| | :: || ||||:| | ||| || ||| | |
Db        264 GDGNFQKWASVNVPRGEEQKYVCVVEHEGLGQPLAVKW 301
```

RESULT 6
Q9BD50
ID Q9BD50 PRELIMINARY; PRT; 340 AA.
AC Q9BD50;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I related protein MR1 isoform.
GN MR1.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RA Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.,
RA Martinez-Naves E.;
RT "Identification of MR1 cDNA sequences in non-human primates.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AJ271828; CAC28215.1; -.
DR HSSP; Q30201; 1A6Z.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 340 AA; 39375 MW; A893952B78725F17 CRC64;

Query Match 35.8%; Score 543.5; DB 7; Length 340;
Best Local Similarity 39.5%; Pred. No. 4.9e-43;
Matches 107; Conservative 51; Mismatches 110; Indels 3; Gaps 3;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||| |      ||  ::: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIRGVPEFISVGYVDSHPITTYDSVTQQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
```

```
Db          83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy         124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
               | |  ||| | | :  |   |  || ::   :  : :|| :| | |:: || |:  | :
Db         142 LIFNKDTLSWLAVDNVAHTIKRAWEANQHELQYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy         184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   |||| |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db         202 EPPLVRVNRKETFPGVTTLFCKAHGFYPPEIYMTWMKNGEEI-VQEMDYGDILPSGDGTY 260

Qy         243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
               | | :  : |     |:| ||| |:   | |
Db         261 QTWASFELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 7
Q9TQK3
ID   Q9TQK3      PRELIMINARY;      PRT;   334 AA.
AC   Q9TQK3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I-related protein MR1 (Fragment).
GN   MR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=99003494; PubMed=9784382;
RA   Yamaguchi H., Kurosawa Y., Hashimoto K.;
RT   "Expanded genomic organization of conserved mammalian MHC class I-
RT   related genes, human MR1 and its murine ortholog.";
RL   Biochem. Biophys. Res. Commun. 250:558-564(1998).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF073485; AAC72900.1; -.
DR   EMBL; AF073484; AAC72900.1; JOINED.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   NON_TER      1      1
SQ   SEQUENCE   334 AA;  38586 MW;  4C3E3A8248A39BA4 CRC64;

  Query Match            35.7%;  Score 542.5;  DB 7;  Length 334;
  Best Local Similarity  39.5%;  Pred. No. 6e-43;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy           4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
               |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db          16 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 75

Qy          64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
               : :| |:||  || |:   :  ::||| |||  | ::|||: || || |: :| |||||
Db          76 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 134

Qy         124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
               | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db         135 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 194

Qy         184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db         195 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 253

Qy         243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
               | | :: : |     |:| ||| |:   | |
Db         254 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 284


RESULT 8
Q9NPL2
ID   Q9NPL2      PRELIMINARY;      PRT;   341 AA.
AC   Q9NPL2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mr1 protein.
GN   MR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood;
RX   MEDLINE=20470599; PubMed=11019920;
RA   Parra-Cuadrado J.F., Navarro P. ., Mirones I., Setien F., Oteo M.,
RA   Martinez-Naves E.;
RT   "A study on the polymorphism of human MHC class I-related MR1 gene and
RT   identification of an MR1-like pseudogene.";
RL   Tissue Antigens 56:170-172(2000).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ249778; CAB77667.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   341 AA;  39366 MW;  2990C1F3F0A1CAD9 CRC64;

  Query Match            35.7%;  Score 542.5;  DB 4;  Length 341;
  Best Local Similarity  39.5%;  Pred. No. 6.2e-43;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy           4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
               |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db          23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy          64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
               : :| |:||  || |:   :  ::||| |||  | ::|||: || || |: :| |||||
Db          83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy         124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
               | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db         142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy         184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db         202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy         243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
               | | :: : |     |:| ||| |:   | |
```

```
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 9
Q95460
ID   Q95460      PRELIMINARY;      PRT;   341 AA.
AC   Q95460;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Class I histocompatibility antigen-like protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=95350662; PubMed=7624800;
RA   Hashimoto K., Hirai M., Kurosawa Y.;
RT   "A gene outside the human MHC related to classical HLA class I
RT   genes.";
RL   Science 269:693-695(1995).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; U22963; AAC50174.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   Genew; HGNC:4975; HLALS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   341 AA;  39366 MW;  2990C1F3F0A1CAD9 CRC64;

  Query Match             35.7%;  Score 542.5;  DB 7;  Length 341;
  Best Local Similarity   39.5%;  Pred. No. 6.2e-43;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::||| ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   | || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|    |   || | |:|  :|| | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 10
Q9BCU3
ID   Q9BCU3      PRELIMINARY;      PRT;   341 AA.
AC   Q9BCU3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I related protein, MR1B1 isoform.
GN   MR1.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martinez-Naves E.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ275984; CAC34272.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   VARIANT     197    197       I -> T (IN REF. 2).
SQ   SEQUENCE   341 AA;  39394 MW;  FBFA822BCAB2C7A8 CRC64;

  Query Match             35.7%;  Score 542.5;  DB 7;  Length 341;
  Best Local Similarity   39.5%;  Pred. No. 6.2e-43;
  Matches  107;  Conservative   51;  Mismatches  110;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::||| ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   | || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDILQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|    |   || | |:|  :|| | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QTWASVELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 11
Q95HB3
ID   Q95HB3      PRELIMINARY;      PRT;   354 AA.
AC   Q95HB3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I antigen alpha chain (Fragment).
OS   Anas platyrhynchos (Domestic duck).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX   NCBI_TaxID=8839;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF393511; AAK84356.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; MHC; Transmembrane.
FT   NON_TER       1       1
SQ   SEQUENCE   354 AA;  39405 MW;  D44FACEE7B60468F CRC64;

  Query Match            35.6%;  Score 540.5;  DB 7;  Length 354;
  Best Local Similarity  37.3%;  Pred. No. 1e-42;
  Matches  101;  Conservative   60;  Mismatches  109;  Indels    1;  Gaps    1;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQL 65
              ||| |      |:   |:  | |:|||| ::|| || |:| :|||  |::  :  | |
Db         24 HSLRYFATAVSDPSPGVPQFVAVGYVDGEVFVRYDSETRGMEPRVDWIADNMDQQYWNGE 83

Qy         66 SQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLE 125
              :::|:| : :: ||  |: | :| |: ||||| : ||:: || |  :|:::|||:|:: :
Db         84 TENLRGAEQIYRVDLETLRERYNQSRGSHTLQHMFGCDLLEDRSISGFFQYGYEGREFIA 143

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
                ||  : ||:  |  || :||      | : : |||  |   |:: :  |: ||:::
Db        144 LDKDTWTFTAADAAAQITKRKWEEDGTVAERRKYYLENTCIEWLRKYVRYGKDVLERRER 203

Qy        186 PLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGW 245
              | |:|:      : :| |||  :||: |:: ||||   :  :| :    :|| |||  |
Db        204 PEVRVSGMEADKILSLSCRAHGFYPRPISISWLKDGM-VQEQETQRGSTVPNSDGTYHIW 262

Qy        246 ITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                |: | ||:: :| |:|||  | || :  ||
Db        263 ATIDVVPGDKDKYQCRVEHASLPQPGLFSWE 293


RESULT 12
Q9BCU4
ID   Q9BCU4      PRELIMINARY;      PRT;   341 AA.
AC   Q9BCU4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I related protein, MR1B1 isoform.
GN   MR1.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martinez-Naves E.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ275982; CAC34274.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   VARIANT     197    197       T -> I (IN REF. 2).
SQ   SEQUENCE   341 AA;  39382 MW;  DFF16AF1FAB2D272 CRC64;

  Query Match            35.5%;  Score 539.5;  DB 7;  Length 341;
  Best Local Similarity  39.5%;  Pred. No. 1.2e-42;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||       ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QTWASVELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 13
Q9HC71
ID   Q9HC71      PRELIMINARY;      PRT;   105 AA.
AC   Q9HC71;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemochromatosis splice variant 838-2283del (Fragment).
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
DR   EMBL; AF144239; AAG29574.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   105 AA;  12233 MW;  4A50E52AA275D4B0 CRC64;

  Query Match            34.9%;  Score 530;  DB 4;  Length 105;
  Best Local Similarity  94.0%;  Pred. No. 2.1e-42;
```

```
  Matches   94;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         89 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 148
              |:||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db          1 HTKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRASEPRAWPTKLEWE 60

Qy        149 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLV 188
              ||:||||| |||||||||||||||||||||||||||| |
Db         61 GHKVRARQNGAYLERDCPAQLQQLLELGRGVLDQQVPPKV 100


RESULT 14
Q8HX66
ID   Q8HX66      PRELIMINARY;      PRT;   356 AA.
AC   Q8HX66;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I antigen (Fragment).
GN   SLA-1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martens G.W., Baker J.E., Smith D.M.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY135589; AAN35107.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   356 AA;  39585 MW;  94FC7A461DBF555B CRC64;

  Query Match             34.3%;  Score 521;  DB 7;  Length 356;
  Best Local Similarity   39.9%;  Pred. No. 7e-41;
  Matches  110;  Conservative   48;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMWL 63
              ||| | :   |  ||| | | |:|||||  || :| :  : |:||| ||:      : |
Db         19 HSLRYFYTAVSRPDLGDSRFIAVGYVDDTQFVRFDSDAPNPRMEPRAPWIQQE-GQEYWD 77

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDN-STEGYWKYGYDGQ 121
              : :::  |      ||  |:   :| |:  |||:| : ||::  |     || :  |||
Db         78 EETRNAMGSAQNDRVDLKTLRGYYNQSEAGSHTIQRMYGCDVGPDGLLLRGYDQDAYDGA 137

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||   : | | |:|||  |    ||: ||:|:  |
Db        138 DYIALNEDLRSWTAADTAAQITKRKWEAANV-AEQERSYLEGTCVEWLQKYLEMGKDTLQ 196

Qy        182 QQVPPLVKVTHHVTSSV-TTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :  ||   || | :| :  |||| || :||: |:: | :: |   ::: |  :  |:|||
Db        197 RAEPPKTHVTRHPSSDLGVTLRCWALGFYPKEISLTWQREGQD-QSQDMELVETRPSGDG 255

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| |  | ||||||| ||| |:| || :|| : |:
Db        256 TFQKWAALVVPPGEEQSYTCHVQHEGLQEPLTLRWD 291


RESULT 15
Q30990
ID   Q30990      PRELIMINARY;      PRT;   332 AA.
AC   Q30990;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chimpanzee MHC class I ChlA chain (Fragment).
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89235215; PubMed=2715640;
RA   Parham P., Lawlor D.A., Lomen C.E., Ennis P.D.;
RT   "Diversity and diversification of HLA-A,B,C alleles.";
RL   J. Immunol. 142:3937-3950(1989).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; M24047; AAA35426.1; -.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   NON_TER     332    332
SQ   SEQUENCE   332 AA;  37433 MW;  9AA9A55DF9E79360 CRC64;

  Query Match             34.2%;  Score 520;  DB 7;  Length 332;
  Best Local Similarity   40.1%;  Pred. No. 7.9e-41;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::| |        ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: August  5, 2003, 13:10:00
Job time : 34 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:10:04 ; Search time 33 Seconds
(without alignments)
993.264 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1520	100.0	276	15	US-10-092-404-2	Sequence 2, Appli
2	1513	99.5	276	15	US-10-092-404-1	Sequence 1, Appli
3	1513	99.5	348	12	US-09-981-606-2	Sequence 2, Appli
4	1493	98.2	276	15	US-10-092-404-3	Sequence 3, Appli
5	514	33.8	92	14	US-10-016-634A-120	Sequence 120, App
6	506	33.3	280	15	US-10-073-300-6	Sequence 6, Appli
7	506	33.3	415	15	US-10-073-300-5	Sequence 5, Appli
8	492	32.4	298	15	US-10-205-823-40	Sequence 40, Appl
9	492	32.4	298	15	US-10-205-823-42	Sequence 42, Appl
10	492	32.4	298	15	US-10-177-293-23	Sequence 23, Appl
11	477	31.4	542	15	US-10-015-535-32	Sequence 32, Appl
12	477	31.4	542	15	US-10-015-535-34	Sequence 34, Appl
13	475	31.2	542	15	US-10-015-535-36	Sequence 36, Appl
14	474	31.2	540	15	US-10-015-535-22	Sequence 22, Appl
15	474	31.2	541	15	US-10-015-535-28	Sequence 28, Appl
16	474	31.2	542	15	US-10-015-535-24	Sequence 24, Appl
17	474	31.2	542	15	US-10-015-535-26	Sequence 26, Appl
18	448	29.5	332	9	US-09-870-521-3	Sequence 3, Appli
19	445	29.3	540	15	US-10-015-535-30	Sequence 30, Appl
20	444	29.2	334	9	US-09-870-521-4	Sequence 4, Appli
21	358.5	23.6	170	9	US-09-925-301-1307	Sequence 1307, Ap
22	336	22.1	271	9	US-09-925-301-1431	Sequence 1431, Ap
23	279	18.4	181	11	US-09-013-077A-13	Sequence 13, Appl
24	275	18.1	145	9	US-09-810-560-8	Sequence 8, Appli
25	243	16.0	184	10	US-09-858-580-21	Sequence 21, Appl
26	243	16.0	184	11	US-09-847-172-21	Sequence 21, Appl
27	226	14.9	91	9	US-09-864-761-38005	Sequence 38005, A
28	223.5	14.7	171	15	US-10-144-929-116	Sequence 116, App
29	223	14.7	91	9	US-09-864-761-35461	Sequence 35461, A
30	210.5	13.8	104	9	US-09-925-302-835	Sequence 835, App
31	208.5	13.7	183	15	US-10-036-542-62	Sequence 62, Appl
32	207	13.6	117	9	US-09-810-560-9	Sequence 9, Appli
33	196.5	12.9	93	9	US-09-864-761-39479	Sequence 39479, A
34	196.5	12.9	110	9	US-09-864-761-35339	Sequence 35339, A
35	196.5	12.9	114	9	US-09-864-761-37988	Sequence 37988, A
36	174.5	11.5	261	10	US-09-925-664-30	Sequence 30, Appl
37	174	11.4	411	14	US-10-015-536-17	Sequence 17, Appl
38	173	11.4	110	10	US-09-796-692-799	Sequence 799, App
39	173	11.4	110	10	US-09-796-692-2139	Sequence 2139, Ap
40	173	11.4	110	15	US-10-040-862-799	Sequence 799, App
41	173	11.4	110	15	US-10-040-862-2139	Sequence 2139, Ap
42	171.5	11.3	285	10	US-09-756-983-24	Sequence 24, Appl
43	167	11.0	772	9	US-09-815-837-74	Sequence 74, Appl
44	166.5	11.0	448	14	US-10-081-281-111	Sequence 111, App
45	166	10.9	246	9	US-09-989-722-225	Sequence 225, App

ALIGNMENTS

RESULT 1
US-10-092-404-2
; Sequence 2, Application US/10092404
; Publication No. US20030073627A1
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; Bjorkman, Pamela J.
; Schatzman, Randall C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
; AND IRON DEFICIENCY DISEASES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/092,404
; FILING DATE: 04-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,964
; FILING DATE: June 12, 1998
; APPLICATION NUMBER: 08/876,010
; FILING DATE: June 13, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0074-999
; TELECOMMUNICATION INFORMATION:

```
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-092-404-2

  Query Match             100.0%;  Score 1520;  DB 15;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-146;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-10-092-404-1
; Sequence 1, Application US/10092404
; Publication No. US20030073627A1
;   GENERAL INFORMATION:
;        APPLICANT: Feder, John N.
;                   Bjorkman, Pamela J.
;                   Schatzman, Randall C.
;        TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
;                            DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;                            AND IRON DEFICIENCY DISEASES
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Pennie & Edmonds, LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036-2811
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/092,404
;             FILING DATE: 04-Mar-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/094,964
;             FILING DATE: June 12, 1998
;             APPLICATION NUMBER: 08/876,010
;             FILING DATE: June 13, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Poissant, Brian M
;             REGISTRATION NUMBER: 28,462
;             REFERENCE/DOCKET NUMBER: 8907-0074-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-092-404-1

  Query Match             99.5%;  Score 1513;  DB 15;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 8.6e-146;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 3
US-09-981-606-2
; Sequence 2, Application US/09981606
; Publication No. US20030129595A1
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg et al.
;  TITLE OF INVENTION: Mutations associated with iron disorders
;  FILE REFERENCE: 24065-004CON
;  CURRENT APPLICATION NUMBER: US/09/981,606
;  CURRENT FILING DATE:  2002-10-16
;  PRIOR APPLICATION NUMBER: 09/277,457
;  PRIOR FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-981-606-2

  Query Match             99.5%;  Score 1513;  DB 12;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-145;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142
```

```
Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
US-10-092-404-3
; Sequence 3, Application US/10092404
; Publication No. US20030073627A1
;   GENERAL INFORMATION:
;        APPLICANT: Feder, John N.
;                   Bjorkman, Pamela J.
;                   Schatzman, Randall C.
;        TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
;                            DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;                            AND IRON DEFICIENCY DISEASES
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Pennie & Edmonds, LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036-2811
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/092,404
;             FILING DATE: 04-Mar-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/094,964
;             FILING DATE: June 12, 1998
;             APPLICATION NUMBER: 08/876,010
;             FILING DATE: June 13, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Poissant, Brian M
;             REGISTRATION NUMBER: 28,462
;             REFERENCE/DOCKET NUMBER: 8907-0074-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-092-404-3

  Query Match             98.2%;  Score 1493;  DB 15;  Length 276;
  Best Local Similarity   98.9%;  Pred. No. 9.3e-144;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 5
US-10-016-634A-120
; Sequence 120, Application US/10016634A
; Publication No. US20020192666A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Ghosh, Malavika
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pro
;  FILE REFERENCE: DEX-0255
;  CURRENT APPLICATION NUMBER: US/10/016,634A
;  CURRENT FILING DATE:  2001-10-31
;  PRIOR APPLICATION NUMBER: US 60/244,258
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 176
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
;   LENGTH: 92
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-016-634A-120

  Query Match             33.8%;  Score 514;  DB 14;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-45;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 60

Qy        152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              ||||||||||||||||||||||||||||||||
Db         61 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 92


RESULT 6
US-10-073-300-6
; Sequence 6, Application US/10073300
; Publication No. US20030003535A1
; GENERAL INFORMATION:
;  APPLICANT: Reiter, Yoram
;  TITLE OF INVENTION: SINGLE CHAIN CLASS I MAJOR HISTO- COMPATIBILITY COMPLEXES
;  FILE REFERENCE: 02/23339
;  CURRENT APPLICATION NUMBER: US/10/073,300
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-073-300-6

  Query Match             33.3%;  Score 506;  DB 15;  Length 280;
  Best Local Similarity   39.4%;  Pred. No. 2.7e-43;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;
```

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |  .  |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db          2 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 60

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db         61 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        121 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 179

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :    |  :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        180 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 238

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        239 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 275


RESULT 7
US-10-073-300-5
; Sequence 5, Application US/10073300
; Publication No. US20030003535A1
; GENERAL INFORMATION:
;  APPLICANT: Reiter, Yoram
;  TITLE OF INVENTION: SINGLE CHAIN CLASS I MAJOR HISTO- COMPATIBILITY COMPLEXES
;  FILE REFERENCE: 02/23339
;  CURRENT APPLICATION NUMBER: US/10/073,300
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human beta2 microglobulin linked to MHC class I heavy chain
US-10-073-300-5

  Query Match             33.3%;  Score 506;  DB 15;  Length 415;
  Best Local Similarity   39.4%;  Pred. No. 4.7e-43;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |  .  |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db        117 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 175

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db        176 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 235

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        236 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 294

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :    |  :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        295 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 353

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        354 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 390


RESULT 8
US-10-205-823-40
; Sequence 40, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Anderson, Dustin
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF PROSTATE CANCER
;  FILE REFERENCE: MRI-044
;  CURRENT APPLICATION NUMBER: US/10/205,823
;  CURRENT FILING DATE:  2002-07-25
;  PRIOR APPLICATION NUMBER: 60/307,982
;  PRIOR FILING DATE: 2001-07-25
;  PRIOR APPLICATION NUMBER: 60/314,356
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/325,020
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: 60/341,746
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-205-823-40

  Query Match             32.4%;  Score 492;  DB 15;  Length 298;
  Best Local Similarity   36.7%;  Pred. No. 7.9e-42;
  Matches  101;  Conservative   53;  Mismatches  111;  Indels   10;  Gaps    4;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQL 65
              :|| |:: | |:    :  |:||| ::|  |  |: : |: :|   |       : | |
Db         28 YSLTYIYTGLSKHVEDVPAFQALGSLNDLQFFRYNSKDRKSQPMGLWRQVE-GMEDWKQD 86

Qy         66 SQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLE 125
              ||  |  : :|       |:| :| |  || ||   |||:: : |:  :||| |||:|::|
Db         87 SQLQKAREDIFMETLKDIVEYYNDSNGSHVLQGRFGCEIENNRSSGAFWKYYYDGKDYIE 146

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              |  :   |   :| |  || :||   :  :: :|||| :||| |:: |:  : :||:| |
Db        147 FNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKYLKYSKNILDRQDP 206

Qy        186 PLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPK---DVLPNGDGT 241
              | | || |         |:| | ::||  | : | :  :       ||:   ||| ||:||
Db        207 PSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EPELRGDVLHNGNGT 261

Qy        242 YQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              || |: :|||| :   |:| |:|  | |||:| ||
Db        262 YQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 296


RESULT 9
US-10-205-823-42
; Sequence 42, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
```

; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-42

Query Match 32.4%; Score 492; DB 15; Length 298;
Best Local Similarity 36.7%; Pred. No. 7.9e-42;
Matches 101; Conservative 53; Mismatches 111; Indels 10; Gaps 4;

```
Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQL 65
              :|| |:: | |:    :  |:||| ::|  |  |: : |: :|   |       : | |
Db         28 YSLTYIYTGLSKHVEDVPAFQALGSLNDLQFFRYNSKDRKSQPMGLWRQVE-GMEDWKQD 86

Qy         66 SQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLE 125
              ||  |  : :|      |:| :| |  || ||   |||:: : |:  :||| |||:|::|
Db         87 SQLQKAREDIFMETLKDIVEYYNDSNGSHVLQGRFGCEIENNRSSGAFWKYYYDGKDYIE 146

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              |  :   |   :| | || :||   :  :: :|||| :||| |:: |:  : :||:| |
Db        147 FNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKYLKYSKNILDRQDP 206

Qy        186 PLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPK---DVLPNGDGT 241
              | | || |      |:| | ::|| | : | :  :        ||:   ||| ||:||
Db        207 PSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EPELRGDVLHNGNGT 261

Qy        242 YQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              || |: :|||| :   |:| |:|  | |||:| ||
Db        262 YQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 296
```

RESULT 10
US-10-177-293-23
; Sequence 23, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-23

Query Match 32.4%; Score 492; DB 15; Length 298;
Best Local Similarity 36.7%; Pred. No. 7.9e-42;
Matches 101; Conservative 53; Mismatches 111; Indels 10; Gaps 4;

```
Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQL 65
              :|| |:: | |:    :  |:||| ::|  |  |: : |: :|   |       : | |
Db         28 YSLTYIYTGLSKHVEDVPAFQALGSLNDLQFFRYNSKDRKSQPMGLWRQVE-GMEDWKQD 86

Qy         66 SQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLE 125
              ||  |  : :|      |:| :| |  || ||   |||:: : |:  :||| |||:|::|
Db         87 SQLQKAREDIFMETLKDIVEYYNDSNGSHVLQGRFGCEIENNRSSGAFWKYYYDGKDYIE 146

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              |  :   |   :| | || :||   :  :: :|||| :||| |:: |:  : :||:| |
Db        147 FNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKYLKYSKNILDRQDP 206

Qy        186 PLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPK---DVLPNGDGT 241
              | | || |      |:| | ::|| | : | :  :        ||:   ||| ||:||
Db        207 PSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EPELRGDVLHNGNGT 261

Qy        242 YQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              || |: :|||| :   |:| |:|  | |||:| ||
Db        262 YQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 296
```

RESULT 11
US-10-015-535-32
; Sequence 32, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
; APPLICANT: Kranz, David M.
; APPLICANT: Brophy, Susan
; TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
; TITLE OF INVENTION: Complexes
; FILE REFERENCE: 100-00
; CURRENT APPLICATION NUMBER: US/10/015,535
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/254,495
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

```
;   OTHER INFORMATION: peptide
US-10-015-535-32

  Query Match             31.4%;  Score 477;  DB 15;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 6e-40;
  Matches  109;  Conservative   41;  Mismatches  118;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDD--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 12
US-10-015-535-34
; Sequence 34, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-34

  Query Match             31.4%;  Score 477;  DB 15;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 6e-40;
  Matches  109;  Conservative   41;  Mismatches  118;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDD--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 13
US-10-015-535-36
; Sequence 36, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-36

  Query Match             31.2%;  Score 475;  DB 15;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 9.5e-40;
  Matches  109;  Conservative   41;  Mismatches  118;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFY--DDESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :  | |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFNSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 14
US-10-015-535-22
; Sequence 22, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 540
```

```
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-22

  Query Match             31.2%;  Score 474;  DB 15;  Length 540;
  Best Local Similarity   39.5%;  Pred. No. 1.2e-39;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDD--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db        144 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 202

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        203 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 262

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db        263 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 321

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        322 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 380

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        381 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 416


RESULT 15
US-10-015-535-28
; Sequence 28, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 541
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-28

  Query Match             31.2%;  Score 474;  DB 15;  Length 541;
  Best Local Similarity   39.5%;  Pred. No. 1.2e-39;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDD--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


Search completed: August  5, 2003, 13:21:55
Job time : 34 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:06:39 ; Search time 15 Seconds
(without alignments)
1769.504 Million cell updates/sec

Title: US-10-092-404-3
Perfect score: 1514
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1129	74.6	359	2	JC5382	hereditary hemochr
2	530.5	35.0	341	2	A57136	class I histocompa
3	517	34.1	361	1	HLRB	MHC class I histoc
4	517	34.1	361	2	I46858	MHC class I RLA pr
5	514	33.9	332	2	S06424	MHC class I histoc
6	511	33.8	365	2	I36961	MHC class I protei
7	510	33.7	361	2	B27638	MHC class I histoc
8	509	33.6	365	2	I83063	A11.2 - human
9	508	33.6	365	2	A47636	MHC class I histoc
10	508	33.6	365	2	I56039	HLA-A30.3 precurso
11	506	33.4	370	1	HLHUA3	MHC class I histoc
12	504	33.3	365	2	I38439	MHC class I histoc
13	503	33.2	365	2	I37542	MHC class I histoc
14	503	33.2	365	2	I38442	gene HLA-A-0205 pr
15	503	33.2	365	2	I61902	MHC class I histoc
16	502	33.2	365	2	I72170	MHC class I histoc
17	502	33.2	365	2	I38441	gene HLA-A-6802 pr
18	501	33.1	355	2	T28149	MHC class I histoc
19	500	33.0	365	1	HLHUA2	MHC class I histoc
20	500	33.0	365	2	I37482	MHC class I histoc
21	500	33.0	365	2	I38519	MHC class I histoc
22	500	33.0	365	2	I84448	MHC class I histoc
23	499	33.0	365	2	I38610	MHC class I histoc
24	499	33.0	365	2	I37470	HLA-A*0210 - human
25	498	32.9	364	2	S03535	class I histocompa
26	497	32.8	365	2	I37476	MHC class I histoc
27	497	32.8	365	2	I37478	MHC class I histoc
28	497	32.8	365	2	I38443	gene HLA-A-0203 pr
29	497	32.8	365	2	I61857	MHC HLA-A2.4a chai
30	496	32.8	357	2	I36965	MHC class I protei
31	495.5	32.7	362	2	A45845	MHC class I histoc
32	495	32.7	365	2	I61856	MHC class I histoc
33	495	32.7	365	2	I54493	MHC class I histoc
34	494	32.6	273	1	HLHU69	MHC class I histoc
35	494	32.6	279	2	JX0353	zinc-alpha 2-glyco
36	494	32.6	365	2	S77963	MHC class I histoc
37	494	32.6	365	2	S01171	class I histocompa
38	494	32.6	365	2	I54416	HLA-AW24 protein -
39	493	32.6	365	2	I37483	HLA-Aw34.2 antigen
40	492	32.5	273	1	HLHUAW	MHC class I histoc
41	492	32.5	360	2	A27638	MHC class I histoc
42	492	32.5	365	2	I72171	HLA-Aw33.1, HLA-Aw
43	491.5	32.5	339	2	I56071	MHC class I histoc
44	491	32.4	362	2	I68724	MHC class I histoc
45	490.5	32.4	341	2	JC5663	major histocompati

ALIGNMENTS

RESULT 1
JC5382
hereditary hemochromatosis protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C;Accession: JC5382
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Biochem. Biophys. Res. Commun. 230, 35-39, 1997
A;Title: Identification of a mouse homolog for the human hereditary haemochromatosis ca
A;Reference number: JC5382; MUID:97148566; PMID:9020055
A;Accession: JC5382
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-359 <HAS>
A;Cross-references: GB:U66849; NID:g1519484; PIDN:AAB07525.1; PID:g1519485
C;Comment: This protein plays a role in iron metabolism.
C;Genetics:
A;Gene: mr2
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-359/Product: hereditary haemochromatosis protein #status predicted <MAT>
F;30-117/Domain: alpha 1 #status predicted <ALF1>
F;118-217/Domain: alpha 2 #status predicted <ALF2>
F;218-309/Domain: alpha 3 #status predicted <ALF3>
F;314-340/Domain: transmembrane #status predicted <TRM>
F;341-359/Domain: intracellular #status predicted <INT>

Query Match 74.6%; Score 1129; DB 2; Length 359;
Best Local Similarity 71.9%; Pred. No. 1.9e-86;
Matches 202; Conservative 30; Mismatches 41; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |:||||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK--------ESHTLQVILGCEMQEDNSTEGYWK 115
               |||||||||:|| ||||||| |:| ||        ||| |||:|||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              ||||||| ||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:||||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310
```

```
RESULT 2
A57136
class I histocompatibility antigen related protein MR1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 23-Jul-1999
C;Accession: A57136
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Science 269, 693-695, 1995
A;Title: A gene outside the human MHC related to classical HLA class I genes.
A;Reference number: A57136; MUID:95350662; PMID:7624800
A;Accession: A57136
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-341 <HAS>
A;Cross-references: GB:U22963; NID:g940353; PIDN:AAC50174.1; PID:g940354
C;Genetics:
A;Gene: GDB:HLALS
A;Cross-references: GDB:683188; OMIM:600764
A;Map position: 1q25.3-1q25.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

  Query Match             35.0%;  Score 530.5;  DB 2;  Length 341;
  Best Local Similarity   39.1%;  Pred. No. 1.2e-36;
  Matches  106;  Conservative   50;  Mismatches  112;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | |  :   |   |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 3
HLRB
MHC class I histocompatibility antigen RLA alpha chain precursor (RL-5) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 22-Jun-1999
C;Accession: A02193
R;Tykocinski, M.L.; Marche, P.N.; Max, E.E.; Kindt, T.J.
J. Immunol. 133, 2261-2269, 1984
A;Title: Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expr
A;Reference number: A02193; MUID:84290724; PMID:6432910
A;Accession: A02193
A;Molecule type: mRNA
A;Residues: 1-361 <TYK>
A;Cross-references: GB:K02441; NID:g1293894; PIDN:AAA98729.1; PID:g165496
A;Note: the source of this protein is a T-lymphoid cell line (RL-5), which has been tran
C;Comment: In contrast to the many antigens expressed in mouse (K, D, and L) and human (
 MHC may therefore differ from the HLA and H-2 loci in having limited complexity.
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: class I histocompatibility antigen RLA alpha chain #status predicted <
F;25-307/Domain: extracellular #status predicted <EXT>
F;25-114/Domain: alpha-1 <EX1>
F;115-206/Domain: alpha-2 <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;308-329/Domain: transmembrane #status predicted <TMM>
F;330-361/Domain: intracellular #status predicted <INT>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

  Query Match             34.1%;  Score 517;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 1.7e-35;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              | :    |   | ||:  |  |: :||      | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 4
I46858
MHC class I RLA precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 21-Jan-2000
C;Accession: I46858
R;Marche, P.N.; Tykocinski, M.L.; Max, E.E.; Kindt, T.J.
Immunogenetics 21, 71-82, 1985
A;Title: Structure of a functional rabbit class I MHC gene: Similarity to human class I
A;Reference number: I46858; MUID:85103547; PMID:3917974
A;Accession: I46858
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-361 <MAR>
A;Cross-references: GB:K02819; NID:g165497; PIDN:AAA98730.1; PID:g165498
C;Genetics:
A;Introns: 25/1; 115/1; 207/1; 299/1; 337/1; 348/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             34.1%;  Score 517;  DB 2;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 1.7e-35;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              | :    |   | ||:  |  |: :||      | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 5
S06424
MHC class I histocompatibility antigen Ch25 alpha chain precursor - chimpanzee
N;Alternate names: MHC Ch1A chain
C;Species: Pan troglodytes (chimpanzee)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999
C;Accession: S06424; I36959
R;Lawlor, D.A.; Ward, F.E.; Ennis, P.D.; Jackson, A.P.; Parham, P.
Nature 335, 268-271, 1988
A;Title: HLA-A and B polymorphisms predate the divergence of humans and chimpanzees.
A;Reference number: S06424; MUID:88319000; PMID:3412487
A;Accession: S06424
A;Molecule type: mRNA
A;Residues: 1-332 <LAW>
R;Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.
J. Immunol. 142, 3937-3950, 1989
A;Title: Diversity and diversification of HLA-A,B,C alleles.
A;Reference number: I36956; MUID:89235215; PMID:2715640
A;Accession: I36959
A;Molecule type: mRNA
A;Residues: 1-332 <RES>
A;Cross-references: GB:M24047; NID:g176818; PIDN:AAA35426.1; PID:g553155
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: glycoprotein; membrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-114/Domain: alpha-1 #status predicted <EX1>
F;115-206/Domain: alpha-2 #status predicted <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;307-331/Domain: transmembrane #status predicted <TMM>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

Query Match 33.9%; Score 514; DB 2; Length 332;
Best Local Similarity 40.1%; Pred. No. 2.7e-35;
Matches 111; Conservative 43; Mismatches 115; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::| |        ||  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||     | | |||||  |    |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :|  :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 6
I36961
MHC class I protein - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I36961
R;Lawlor, D.A.; Warren, E.; Ward, F.E.; Parham, P.
Immunol. Rev. 113, 147-185, 1990
A;Title: Comparison of class I MHC alleles in humans and apes.
A;Reference number: I36961; MUID:90201944; PMID:1690682
A;Accession: I36961
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30678; NID:g176822; PIDN:AAA87970.1; PID:g176823
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 33.8%; Score 511; DB 2; Length 365;
Best Local Similarity 39.7%; Pred. No. 5.4e-35;
Matches 110; Conservative 44; Mismatches 115; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::| |        ||  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||     | | |||||  |    |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :|  :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 7
B27638
MHC class I histocompatibility antigen alpha chain precursor (BL3-7) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 16-Feb-1997
C;Accession: B27638
R;Ennis, P.D.; Jackson, A.P.; Parham, P.
J. Immunol. 141, 642-651, 1988
A;Title: Molecular cloning of bovine class I MHC cDNA.
A;Reference number: A92826; MUID:88258075; PMID:3133413
A;Accession: B27638
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-361 <ENN>
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: MHC class I histocompatibility antigen, BoLA alpha chain (BL3-7) #stat
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 33.7%; Score 510; DB 2; Length 361;
Best Local Similarity 38.9%; Pred. No. 6.4e-35;
Matches 109; Conservative 49; Mismatches 114; Indels 8; Gaps 7;

```
Qy          2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISS 59
              |  |||| | : | |    ||    | |:|||||  || :| :  : | ||| ||:
Db         23 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 81

Qy         60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
              : | : ::  |     :| ||   |:   :| |:  || :| : ||::  |      |:|::|
Db         82 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 141

Qy        118 YDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
              |||:| :    :  | ||:  |  || :||     |   | ||| :|   |:: || |:
Db        142 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 200

Qy        178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              | :  ||   |||| ::    |||| || :||: |:: | :: :    :: |  :  |
Db        201 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 259

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              :||||:| |  | || ||||||||:|:| || :|| : ||
Db        260 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 299
```

RESULT 8

```
I83063
A11.2 - human
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C;Accession: I83063
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.;
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I83063
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16842; NID:g540517; PIDN:BAA04118.1; PID:g487911
C;Genetics:
A;Gene: A1102
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.6%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 7.9e-35;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :    |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGKPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|      ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||: |  || :||     | | |||||  |   |:: || |: |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 9
A47636
MHC class I histocompatibility antigen HLA-A11 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 28-Apr-1995 #text_change 23-Jul-1999
C;Accession: S03536; S03694; A47636; I60129
R;Mayer, W.E.; Jonker, M.; Klein, D.; Ivanyi, P.; van Seventer, G.; Klein, J.
EMBO J. 7, 2765-2774, 1988
A;Title: Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-spec
A;Reference number: S01171; MUID:89030641; PMID:2460344
A;Accession: S03536
A;Molecule type: mRNA
A;Residues: 1-365 <MAY>
A;Cross-references: EMBL:X13111; NID:g32138; PIDN:CAA31503.1; PID:g32139
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
A;Accession: S03694
A;Molecule type: mRNA
A;Residues: 1-42,'K',44-298 <MA2>
A;Cross-references: EMBL:X13112; NID:g32142; PIDN:CAA31504.1; PID:g32143
A;Note: this allele is designated A*1102 (formerly A11K, A11.2)
R;Cowan, E.P.; Jelachich, M.L.; Biddison, W.E.; Coligan, J.E.
Immunogenetics 25, 241-250, 1987
A;Title: DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows identification
A;Reference number: A47636; MUID:87192928; PMID:2437024
A;Accession: A47636
A;Molecule type: DNA
A;Residues: 26-365 <COW>
A;Cross-references: GB:M16007; GB:M16008; GB:M16009; GB:M16010; NID:g184130; PIDN:AAA654
A;Note: the authors translated the codon GAC for residue 89 as Ala, CCG for residue 104
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I60129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16841; NID:g540516; PIDN:BAA04117.1; PID:g487909
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: class I histocompatibility antigen alpha chain #status predicted <MAT>
F;25-298/Domain: extracellular #status predicted <EXT>
F;220-285/Domain: immunoglobulin homology <IMM>
F;299-337/Domain: transmembrane #status predicted <TMM>
F;338-365/Domain: intracellular #status predicted <INT>

  Query Match             33.6%;  Score 508;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 9.5e-35;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :    |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|      ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||: |  || :||     | | |||||  |   |:: || |: |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 10
I56039
HLA-A30.3 precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I56039
R;Kato, K.; Trapani, J.A.; Allopenna, J.; Dupont, B.; Yang, S.Y.
J. Immunol. 143, 3371-3378, 1989
A;Title: Molecular analysis of the serologically defined HLA-Aw19 antigens. A geneticall
A;Reference number: I56039; MUID:90038496; PMID:2478623
A;Accession: I56039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30576; NID:g187646; PIDN:AAA59612.1; PID:g386878
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.6%;  Score 508;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 9.5e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |    |    |   | |:|||||  || :| :  |:|:||| ||:      : |
```

```
Db         26 SHSMRYFFTSVSRPGSGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              |  ::::|        ||  |:   :| |:  |||:|:: ||::  |      || :: |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYEQHAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  |: :||  :  | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITQRKWEAAR-WAEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 11
HLHUA3
MHC class I histocompatibility antigen HLA-A3 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 02-Sep-1997
C;Accession: A02192
R;Strachan, T.; Sodoyer; R.; Damotte, M.; Jordan, B.R.
EMBO J. 3, 887-894, 1984
A;Title: Complete nucleotide sequence of a functional class I HLA gene, HLA-A3: implicat
A;Reference number: A02192; MUID:84207948; PMID:6609814
A;Accession: A02192
A;Molecule type: DNA
A;Residues: 1-370 <STR>
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
A;Introns: 30/1; 120/1; 212/1; 304/1; 343/1; 354/1; 370/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-370/Product: class I histocompatibility antigen HLA-A3 alpha chain #status predicte
F;30-312/Domain: extracellular #status predicted <EXT>
F;30-119/Domain: alpha-1 <EX1>
F;120-211/Domain: alpha-2 <EX2>
F;225-290/Domain: immunoglobulin homology <IMM>
F;313-337/Domain: transmembrane #status predicted <TMM>
F;338-370/Domain: intracellular #status predicted <INT>
F;115/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;232-288/Disulfide bonds: #status predicted

  Query Match             33.4%;  Score 506;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 1.4e-34;
  Matches  110;  Conservative   46;  Mismatches  112;  Indels   10;  Gaps    8;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:|||| ||:      : |
Db         31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|        ||  |:   :| |:  |||:|:: ||::  |      || :  |||
Db         90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              :| :    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db        150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy        180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db        208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||:| |  : || |||||||| |:| || :|| : ||
Db        267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 12
I38439
MHC class I histocompatibility antigen HLA-A*8001 precursor - human
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 21-Jan-2000
C;Accession: I59638; I38439
R;Domena, J.D.; Hildebrand, W.H.; Bias, W.B.; Parham, P.
Tissue Antigens 42, 156-159, 1993
A;Title: A sixth family of HLA-A alleles defined by HLA-A*8001.
A;Reference number: I59638; MUID:94112691; PMID:8284791
A;Accession: I59638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <DOM>
A;Cross-references: GB:L18898; NID:g306853; PIDN:AAA17012.1; PID:g306854
R;Balas, A.; Garcia-Sanchez, F.; Gomez-Reino, F.; Vicario, J.L.
Immunogenetics 39, 452, 1994
A;Title: Characterization of a new and highly distinguishable HLA-A allele in a Spanish
A;Reference number: I38439; MUID:94245293; PMID:8188325
A;Accession: I38439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <BAL>
A;Cross-references: EMBL:U03754; NID:g432407; PIDN:AAC04322.1; PID:g432408
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.3%;  Score 504;  DB 2;  Length 365;
  Best Local Similarity   38.3%;  Pred. No. 2.1e-34;
  Matches  106;  Conservative   52;  Mismatches  111;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              : ::::|        :  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||  : || | ||||| :|   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| ||:||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|:|||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 13
I37542
MHC class I histocompatibility antigen HLA-A2 alpha chain (allele A*0216) precursor - hu
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I37542; S49582
R;Barouch, D.; Krausa, P.; Bodmer, J.; Browning, M.J.; McMichael, A.J.
Immunogenetics 41, 388, 1995
A;Title: Identification of a novel HLA-A2 subtype, HLA-A*0216.
A;Reference number: I37542; MUID:95278976; PMID:7759139
A;Accession: I37542
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:Z46633; NID:g575248; PIDN:CAA86602.1; PID:g575249
```

```
A;Note: submitted to the EMBL Data Library, November 1994
C;Genetics:
A;Gene: hla-A
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.2%;  Score 503;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 2.5e-34;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||   : | | |||||  :|   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGECVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 14
I38442
gene HLA-A-0205 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I38442
R;Holmes, N.; Ennis, P.; Wan, A.M.; Denney, D.W.; Parham, P.
J. Immunol. 139, 936-941, 1987
A;Title: Multiple genetic mechanisms have contributed to the generation of the HLA-A2/A2
A;Reference number: I38441; MUID:87252273; PMID:3496393
A;Accession: I38442
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:U03862; NID:g432436; PIDN:AAA03603.1; PID:g432437
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.2%;  Score 503;  DB 2;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 2.5e-34;
  Matches  110;  Conservative   43;  Mismatches  116;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |     |   | |:|||||  || :| :  |||:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  ||||| : ||::  |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
I61902
MHC class I histocompatibility antigen HLA-A alpha chain precursor - human (isolate A*0
C;Species: Homo sapiens (man)
A;Variety: isolate A*0212
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C;Accession: I61902
R;Belich, M.P.; Madrigal, J.A.; Hildebrand, W.H.; Zemmour, J.; Williams, R.C.; Luz, R.;
Nature 357, 326-329, 1992
A;Title: Unusual HLA-B alleles in two tribes of Brazilian Indians.
A;Reference number: I37120; MUID:92269955; PMID:1317015
A;Accession: I61902
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M84378; NID:g187625; PIDN:AAA59604.1; PID:g187626
A;Experimental source: cell line KRC 033; isolate A*0212
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: MHC class I histocompatibility antigen HLA-A alpha chain #status pred
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.2%;  Score 503;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 2.5e-34;
  Matches  109;  Conservative   44;  Mismatches  116;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: August  5, 2003, 13:10:37
Job time : 16 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:05:29 ; Search time 9.5 Seconds
(without alignments)
1366.250 Million cell updates/sec

Title: US-10-092-404-3
Perfect score: 1514
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1502	99.2	348	1	HFE_HUMAN	Q30201 homo sapien
2	1227	81.0	348	1	HFE_DICSU	Q9gl42 dicerorhinu
3	1225	80.9	348	1	HFE_CERSI	Q9gkz0 ceratotheri
4	1221	80.6	348	1	HFE_RHIUN	Q9gl41 rhinoceros
5	1218	80.4	348	1	HFE_DICBI	Q9gl43 diceros bic
6	1145	75.6	360	1	HFE_RAT	O35799 rattus norv
7	1129	74.6	359	1	HFE_MOUSE	P70387 mus musculu
8	517	34.1	361	1	HA1A_RABIT	P01894 oryctolagus
9	517	34.1	361	1	HA1B_RABIT	P06140 oryctolagus
10	511	33.8	365	1	1A01_PANTR	P16209 pan troglod
11	510	33.7	364	1	HA1B_BOVIN	P13753 bos taurus
12	508	33.6	365	1	1A11_HUMAN	P13746 homo sapien
13	506	33.4	370	1	1A03_HUMAN	P04439 homo sapien
14	504	33.3	365	1	1A80_HUMAN	Q09160 homo sapien
15	502	33.2	365	1	1A31_HUMAN	P16189 homo sapien
16	500	33.0	365	1	1A02_HUMAN	P01892 homo sapien
17	500	33.0	365	1	1A30_HUMAN	P16188 homo sapien
18	500	33.0	365	1	1A74_HUMAN	P30459 homo sapien
19	498	32.9	365	1	1A03_PANTR	P13748 pan troglod
20	497	32.8	365	1	1A33_HUMAN	P16190 homo sapien
21	497	32.8	365	1	1A36_HUMAN	P30455 homo sapien
22	497	32.8	365	1	1A68_HUMAN	P01891 homo sapien
23	495.5	32.7	362	1	HA19_CANFA	P18466 canis famil
24	495	32.7	365	1	1A01_HUMAN	P30443 homo sapien
25	494	32.6	273	1	1A69_HUMAN	P10316 homo sapien
26	494	32.6	296	1	ZA2G_RAT	Q63678 rattus norv
27	494	32.6	365	1	1A04_PANTR	P13749 pan troglod
28	494	32.6	365	1	1A24_HUMAN	P05534 homo sapien
29	492	32.5	360	1	HA1A_BOVIN	P13752 bos taurus
30	491	32.4	362	1	1B45_HUMAN	P30485 homo sapien
31	490	32.4	365	1	1A23_HUMAN	P30447 homo sapien
32	488	32.2	338	1	1B20_HUMAN	P30467 homo sapien
33	487	32.2	363	1	1B04_GORGO	P30382 gorilla gor
34	486	32.1	295	1	ZA2G_HUMAN	P25311 homo sapien
35	486	32.1	322	1	HA10_MOUSE	P01898 mus musculu
36	486	32.1	362	1	1B29_HUMAN	P18463 homo sapien
37	486	32.1	371	1	HA12_RAT	P16391 rattus norv
38	485	32.0	365	1	1A34_HUMAN	P30453 homo sapien
39	485	32.0	365	1	1A66_HUMAN	P30457 homo sapien
40	484	32.0	338	1	HLAG_HUMAN	P17693 homo sapien
41	484	32.0	361	1	1B14_HUMAN	P03989 homo sapien
42	484	32.0	362	1	1B18_HUMAN	P10318 homo sapien
43	484	32.0	366	1	1C02_GORGO	P30385 gorilla gor
44	484	32.0	366	1	1C04_GORGO	P30387 gorilla gor
45	483	31.9	359	1	1B01_PANTR	P13750 pan troglod

ALIGNMENTS

RESULT 1
HFE_HUMAN
ID HFE_HUMAN STANDARD; PRT; 348 AA.
AC Q30201; O75929; O75930; O75931; Q96KU5; Q96KU7; Q96KU8; Q9HC64;
AC Q9HC68; Q9HC70; Q9HC83;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hereditary hemochromatosis protein precursor (HLA-H).
GN HFE OR HLAH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=96331279; PubMed=8696333;
RA Feder J.N., Gnirke A., Thomas W., Tsuchihashi Z., Ruddy D.A.,
RA Basava A., Dormishian F., Domingo R., Ellis M.C. Jr., Fullan A.,
RA Hinton L.M., Jones N.L., Kimmel B.E., Kronmal G.S., Lauer P.,
RA Lee V.K., Loeb D.B., Mapa F.A., McClelland E., Meyer N.C.,
RA Mintier G.A., Moeller N., Moore T., Morikang E., Prass C.E.,
RA Quintana L., Starnes S.M., Schatzman R.C., Brunke K.J.,
RA Drayna D.T., Risch N.J., Bacon B.R., Wolff R.K.;
RT "A novel MHC class I-like gene is mutated in patients with hereditary
RT haemochromatosis.";
RL Nat. Genet. 13:399-409(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Albig W., Burmester N., Bode C., Doenecke D., Drabent B.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97294057; PubMed=9149941;
RA Ruddy D.A., Kronmal G.S., Lee V.K., Mintier G.A., Quintana L.,
RA Domingo R. Jr., Meyer N.C., Irrinki A., McClelland E.E., Fullan A.,
RA Mapa F.A., Moore T., Thomas W., Loeb D.B., Harmon C., Tsuchihashi Z.,
RA Wolff R.K., Schatzman R.C., Feder J.N.;
RT "A 1.1-Mb transcript map of the hereditary hemochromatosis locus.";
RL Genome Res. 7:441-456(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Gasparini P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).
RX MEDLINE=99180629; PubMed=10079302;
RA Rhodes D.A., Trowsdale J.;
RT "Alternate splice variants of the hemochromatosis gene Hfe.";
RL Immunogenetics 49:357-359(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 5; 6 AND 7).
RA Oliva R., Sanchez M.;
RT "Identification of different alternative splicing forms of the HFE
RT gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.

RN [7]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 7; 8; 9 AND 10).
RX MEDLINE=20448010; PubMed=11001625;
RA Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.-Y.,
RA David V., Mosser J.;
RT "The HFE gene undergoes alternate splicing processes.";
RL Blood Cells Mol. Dis. 26:155-162(2000).
RN [8]
RP FUNCTION.
RX MEDLINE=98132614; PubMed=9465039;
RA Feder J.N., Penny D.M., Irrinki A., Lee V.K., Lebron J.A., Watson N.,
RA Tsuchihashi Z., Sigal E., Bjorkman P.J., Schatzman R.C.;
RT "The hemochromatosis gene product complexes with the transferrin
RT receptor and lowers its affinity for ligand binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1472-1477(1998).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=98206473; PubMed=9546397;
RA Lebron J.A., Bennett M.J., Vaughn D.E., Chirino A.J., Snow P.M.,
RA Mintier G.A., Feder J.N., Bjorkman P.J.;
RT "Crystal structure of the hemochromatosis protein HFE and
RT characterization of its interaction with transferrin receptor.";
RL Cell 93:111-123(1998).
RN [10]
RP VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=97260408; PubMed=9106528;
RA Carella M., D'Ambrosio L., Totaro A., Grifa A., Valentino M.A.,
RA Piperno A., Girelli D., Roetto A., Franco B., Gasparini P.,
RA Camaschella C.;
RT "Mutation analysis of the HLA-H gene in Italian hemochromatosis
RT patients.";
RL Am. J. Hum. Genet. 60:828-832(1997).
RN [11]
RP VARIANT HH/PCT TYR-282.
RX MEDLINE=97176837; PubMed=9024376;
RA Roberts A.G., Whatley S.D., Morgan R.R., Worwood M., Elder G.H.;
RT "Increased frequency of the haemochromatosis Cys282Tyr mutation in
RT sporadic porphyria cutanea tarda.";
RL Lancet 349:321-323(1997).
RN [12]
RP VARIANT HH/PCT ASP-63.
RX MEDLINE=98085904; PubMed=9425935;
RA Sampietro M., Piperno A., Lupica L., Arosio C., Vergani A.,
RA Corbetta N., Malosio I., Mattioli M., Fracanzani A.L.,
RA Cappellini M.D., Fiorelli G., Fargion S.;
RT "High prevalence of the His63Asp HFE mutation in Italian patients with
RT porphyria cutanea tarda.";
RL Hepatology 27:181-184(1998).
RN [13]
RP VARIANTS HH/PCT ASP-63 AND TYR-282.
RX MEDLINE=98281650; PubMed=9620340;
RA Bonkovsky H.L., Poh-Fitzpatrick M., Pimstone N., Obando J.,
RA Di Bisceglie A., Tattrie C., Tortorelli K., LeClair P., Mercurio M.G.,
RA Lambrecht R.W.;
RT "Porphyria cutanea tarda, hepatitis C, and HFE gene mutations in North
RT America.";
RL Hepatology 27:1661-1669(1998).
RN [14]
RP VARIANTS HH ASP-63; CYS-65 AND TYR-282.
RX MEDLINE=99211934; PubMed=10194428;
RA Mura C., Raguenes O., Ferec C.;
RT "HFE mutations analysis in 711 hemochromatosis probands: evidence for
RT S65C implication in mild form of hemochromatosis.";
RL Blood 93:2502-2505(1999).
RN [15]
RP VARIANTS HH CYS-65; ARG-93 AND THR-105.
RX MEDLINE=20042794; PubMed=10575540;
RA Barton J.C., Sawada-Hirai R., Rothenberg B.E., Acton R.T.;
RT "Two novel missense mutations of the HFE gene (I105T and G93R) and
RT identification of the S65C mutation in Alabama hemochromatosis
RT probands.";
RL Blood Cells Mol. Dis. 25:147-155(1999).
RN [16]
RP VARIANTS VP ASP-63 AND HIS-127, VARIANT HH MET-330, AND VARIANTS
RP MET-53 AND MET-59.
RX MEDLINE=99330560; PubMed=10401000;
RA de Villiers J.N.P., Hillermann R., Loubser L., Kotze M.J.;
RT "Spectrum of mutations in the HFE gene implicated in haemochromatosis
RT and porphyria.";
RL Hum. Mol. Genet. 8:1517-1522(1999).
RN [17]
RP VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=99140260; PubMed=10094552;
RA Merryweather-Clarke A.T., Simonsen H., Shearman J.D., Pointon J.J.,
RA Norgaard-Pedersen B., Robson K.J.H.;
RT "A retrospective anonymous pilot study in screening newborns for HFE
RT mutations in Scandinavian populations.";
RL Hum. Mutat. 13:154-159(1999).
RN [18]
RP VARIANT HH CYS-65.
RA Fagan E., Payne S.J.;
RT "A novel missense mutation S65C in the HFE gene with a possible role
RT in hereditary haemochromatosis.";
RL Hum. Mutat. 13:507-508(1999).
RN [19]
RP VARIANT LYS-277.
RX MEDLINE=20081073; PubMed=10612845;
RA Bradbury R., Fagan E., Payne S.J.;
RT "Two novel polymorphisms (E277K and V212V) in the haemochromatosis
RT gene HFE.";
RL Hum. Mutat. 15:120-120(2000).
CC -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC affinity for iron-loaded transferrin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q30201-1; Sequence=Displayed;
CC Name=2; Synonyms=delE2;
CC IsoId=Q30201-2; Sequence=VSP_003218;
CC Name=3; Synonyms=del14E4;
CC IsoId=Q30201-3; Sequence=VSP_003225;
CC Name=4; Synonyms=delE214E4;
CC IsoId=Q30201-4; Sequence=VSP_003218, VSP_003225;
CC Name=5;
CC IsoId=Q30201-5; Sequence=VSP_003219;
CC Name=6;
CC IsoId=Q30201-6; Sequence=VSP_003220;
CC Name=7; Synonyms=delE3;
CC IsoId=Q30201-7; Sequence=VSP_003221;
CC Name=8; Synonyms=1043-2283del,intron6ins;
CC IsoId=Q30201-8; Sequence=VSP_003226, VSP_003227;
CC Name=9; Synonyms=delE3-7;
CC IsoId=Q30201-9; Sequence=VSP_003223, VSP_003224;
CC Name=10; Synonyms=562-878del;
CC IsoId=Q30201-10; Sequence=VSP_003222;
CC -!- TISSUE SPECIFICITY: IN ALL TISSUES TESTED EXCEPT BRAIN.
CC -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF HEREDITARY HEMOCHROMATOSIS
CC (HH). HH IS AN AUTOSOMAL RECESSIVE INBORN DISORDER OF IRON
CC METABOLISM, FREQUENT AMONG CAUCASIANS. HH IS CHARACTERIZED BY
CC ABNORMAL INTESTINAL IRON ABSORPTION AND PROGRESSIVE INCREASE OF
CC TOTAL BODY IRON, WHICH RESULTS IN MIDLIFE IN CLINICAL
CC COMPLICATIONS INCLUDING CIRRHOSIS, CARDIOPATHY, DIABETES,
CC ENDOCRINE DYSFUNCTIONS, ARTHROPATHY, AND SUSCEPTIBILITY TO LIVER
CC CANCER. SINCE THE DISEASE COMPLICATIONS CAN BE EFFECTIVELY
CC PREVENTED BY REGULAR PHLEBOTOMIES, EARLY DIAGNOSIS IS MOST
CC IMPORTANT TO PROVIDE A NORMAL LIFE EXPECTANCY TO THE AFFECTED
CC SUBJECTS.
CC -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF PORPHYRIA CUTANEA TARDA
CC (PCT), A DISORDER CHARACTERIZED BY LIGHT-SENSITIVE DERMATITIS AND
CC PRESENCE OF LARGE AMOUNTS OF UROPORPHYRIN IN URINE. IRON OVERLOAD
CC IS OFTEN PRESENT IN ASSOCIATION WITH VARYING DEGREES OF LIVER
CC DAMAGE. PCT IS THE MOST COMMON FORM OF PORPHYRIA WORLDWIDE. IT
CC OCCURS IN TWO FORMS: THE SPORADIC TYPE (PCT TYPE I) AND THE
CC FAMILIAL TYPE (PCT TYPE II), WHICH IS INHERITED IN AN AUTOSOMAL

```
  Query Match             99.2%;  Score 1502;  DB 1;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 2.3e-118;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 2
HFE_DICSU
ID   HFE_DICSU      STANDARD;      PRT;   348 AA.
AC   Q9GL42;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Dicerorhinus sumatrensis (Sumatran rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.
OX   NCBI_TaxID=89632;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007543; AAG23703.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39740 MW;  518BFD357AB83B90 CRC64;

  Query Match             81.0%;  Score 1227;  DB 1;  Length 348;
  Best Local Similarity   81.0%;  Pred. No. 2.2e-95;
  Matches  221;  Conservative   20;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||||:| || ||||||||||:||| |:||||| | |  ||      ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              |||||||||||| |||||||:||| ||||||||||||:|||||| |:||||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |: |||||||||||||||||||||||   ||
Db        266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 3
HFE_CERSI
ID   HFE_CERSI      STANDARD;      PRT;   348 AA.
AC   Q9GKZ0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX   NCBI_TaxID=9807;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007541; AAG23701.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
```

```
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39822 MW;  2523016ECE9FBE91 CRC64;

  Query Match             80.9%;  Score 1225;  DB 1;  Length 348;
  Best Local Similarity   81.3%;  Pred. No. 3.2e-95;
  Matches  222;  Conservative   18;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |:||||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              |||||||||||| |||||||:||| ||||||||||||:|||||| |:|||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:||||||:|||:|||||||:||||::||| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPMDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
               |  ||||||||||||||||||||||||   ||
Db        266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 4
HFE_RHIUN
ID   HFE_RHIUN      STANDARD;      PRT;   348 AA.
AC   Q9GL41;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Rhinoceros unicornis (Greater Indian rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
OX   NCBI_TaxID=9809;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007544; AAG23704.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39743 MW;  F2723D57A327A6B4 CRC64;

  Query Match             80.6%;  Score 1221;  DB 1;  Length 348;
  Best Local Similarity   80.6%;  Pred. No. 7e-95;
  Matches  220;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |:||||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              |||||||||||| |||||||:||| ||||||||||||:|||||| |:| ||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWMYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
               |: ||||||||||||||||||||||||   ||
Db        266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 5
HFE_DICBI
ID   HFE_DICBI      STANDARD;      PRT;   348 AA.
AC   Q9GL43; Q9GK81;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
```

```
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007542; AAG23702.1; -.
DR   EMBL; AF301592; AAG39940.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39818 MW;  4D95E7B01E48FB90 CRC64;

  Query Match             80.4%;  Score 1218;  DB 1;  Length 348;
  Best Local Similarity   80.6%;  Pred. No. 1.2e-94;
  Matches  220;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |:||||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              ||:|||||||||| |||||||:||| |||||||||||||:|||||| |:|||||||||
Db         86 QLTQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||||  |:||||||:|||:|||||||:||||::|:|  |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                |  |||||||||||||||||||||||    ||
Db        266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298
```

```
RESULT 6
HFE_RAT
ID   HFE_RAT        STANDARD;      PRT;   360 AA.
AC   O35799; O35175;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor (RT1-CAFE).
GN   HFE.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Banasch M.W., Schaefer H., Schmidt W.E.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 42-303 FROM N.A.
RC   TISSUE=Small intestine;
RA   Sawada-Hirai R., Rothenberg B.E.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS TO TRANSFERRIN RECEPTOR (TFR) AND REDUCES ITS
CC       AFFINITY FOR IRON-LOADED TRANSFERRIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001517; CAA04799.1; -.
DR   EMBL; AF008587; AAB86597.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    360       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       26    127       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      128    218       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      219    310       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      311    319       CONNECTING PEPTIDE.
FT   TRANSMEM    320    340       POTENTIAL.
FT   DOMAIN      341    360       CYTOPLASMIC TAIL.
FT   DISULFID    137    200       BY SIMILARITY.
FT   DISULFID    238    295       BY SIMILARITY.
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    198    198       R -> K (IN REF. 2).
SQ   SEQUENCE   360 AA;  40988 MW;  CC819834EE3240B3 CRC64;

  Query Match             75.6%;  Score 1145;  DB 1;  Length 360;
  Best Local Similarity   73.2%;  Pred. No. 1.6e-88;
  Matches  205;  Conservative   29;  Mismatches   38;  Indels    8;  Gaps    1;
```

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQ 64
              |||| |||||||: ||||  |||||||||||||| |:||||| ||| ||:  : |||:|||
Db         32 SHSLRYLFMGASKPDLGLPFFEALGYVDDQLFVSYNHESRRAEPRAPWILGQTSSQLWLQ 91

Qy         65 LSQSLKGWDHMFTVDFWTIMENHNASK--------ESHTLQVILGCEMQEDNSTEGYWKY 116
              |||||||||:|| ||||||| |:| ||        ||| ||||||||: ||||| |:|||
Db         92 LSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLRVVPESHILQVILGCEVHEDNSTSGFWKY 151

Qy        117 GYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELG 176
              |||||| ||||| ||:| ||||||| ||:||| |:|||||:| ||:|||| ||:|:|||
Db        152 GYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHRIRARQSRDYLQRDCPQQLKQVLELQ 211

Qy        177 RGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              |||| |||| ||||| |  |: |:|||:|||::|||||:|||| ||:|||:  |::|||
Db        212 RGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQNITMRWLKDSQPLDAKDVNPENVLP 271

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||:|||| |||| |::|||||||||||||   ||
Db        272 NGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPLTATWE 311
```

RESULT 7
HFE_MOUSE
ID HFE_MOUSE STANDARD; PRT; 359 AA.
AC P70387;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hereditary hemochromatosis protein homolog precursor.
GN HFE OR MR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=98060831; PubMed=9396865;
RA Riegert P., Gilfillan S., Nanda I., Schmid M., Bahram S.;
RT "The mouse HFE gene.";
RL Immunogenetics 47:174-177(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Lung;
RA Hashimoto K.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 37-211 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=97148566; PubMed=9020055;
RA Hashimoto K., Hirai M., Kurosawa Y.;
RT "Identification of a mouse homolog for the human hereditary
RT haemochromatosis candidate gene.";
RL Biochem. Biophys. Res. Commun. 230:35-39(1997).
RN [4]
RP SEQUENCE OF 79-359 FROM N.A.
RC STRAIN=129;
RA Albig W., Drabent B., Doenecke D.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC affinity for iron-loaded transferrin (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF007558; AAC03447.1; -.
DR EMBL; U66849; AAB07525.1; -.
DR EMBL; Y12650; CAA73197.1; -.
DR EMBL; U80604; AAB51504.1; -.
DR PIR; JC5382; JC5382.
DR HSSP; Q30201; 1A6Z.
DR MGD; MGI:109191; Hfe.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 359 HEREDITARY HEMOCHROMATOSIS PROTEIN
FT HOMOLOG.
FT DOMAIN 25 126 EXTRACELLULAR ALPHA-1.
FT DOMAIN 127 217 EXTRACELLULAR ALPHA-2.
FT DOMAIN 218 309 EXTRACELLULAR ALPHA-3.
FT DOMAIN 310 318 CONNECTING PEPTIDE.
FT TRANSMEM 319 339 POTENTIAL.
FT DOMAIN 340 359 CYTOPLASMIC TAIL.
FT DISULFID 136 199 BY SIMILARITY.
FT DISULFID 237 294 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 359 AA; 40548 MW; 4BDE6C27F9FF20B4 CRC64;

Query Match 74.6%; Score 1129; DB 1; Length 359;
Best Local Similarity 71.9%; Pred. No. 3.5e-87;
Matches 202; Conservative 30; Mismatches 41; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| |||||||||| |:||||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK--------ESHTLQVILGCEMQEDNSTEGYWK 115
               |||||||||:|| ||||||| |:| ||        ||| |||:|||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              ||||||| ||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:|||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310
```

RESULT 8
HA1A_RABIT
ID HA1A_RABIT STANDARD; PRT; 361 AA.
AC P01894;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RLA class I histocompatibility antigen, alpha chain 11/11 precursor.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84290724; PubMed=6432910;
RA   Tykocinski M.L., Marche P.N., Max E.E., Kindt T.J.;
RT   "Rabbit class I MHC genes: cDNA clones define full-length transcripts
RT   of an expressed gene and a putative pseudogene.";
RL   J. Immunol. 133:2261-2269(1984).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02441; AAA98729.1; -.
DR   PIR; A02193; HLRB.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN 11/11.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    329
FT   DOMAIN      330    361       CYTOPLASMIC.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40447 MW;  580B673323C1AE35 CRC64;

  Query Match             34.1%;  Score 517;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 4.9e-36;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |      | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              | :   |  | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   ||||  |    |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

```
RESULT 9
HA1B_RABIT
ID   HA1B_RABIT     STANDARD;      PRT;   361 AA.
AC   P06140;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 19-1 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85103547; PubMed=3917974;
RA   Marche P.N., Tykocinski M.L., Max E.E., Kindt T.J.;
RT   "Structure of a functional rabbit class I MHC gene: similarity to
RT   human class I genes.";
RL   Immunogenetics 21:71-82(1985).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02819; AAA98730.1; -.
DR   PIR; I46858; I46858.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN 19-1.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    329
FT   DOMAIN      330    361       CYTOPLASMIC.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40455 MW;  C06FBD8B87ED0546 CRC64;

  Query Match             34.1%;  Score 517;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 4.9e-36;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |      | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144
```

```
Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 10
1A01_PANTR
ID   1A01_PANTR     STANDARD;      PRT;   365 AA.
AC   P16209;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   CHLA class I histocompatibility antigen, A-2 alpha chain precursor.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90201944; PubMed=1690682;
RA   Lawlor D.A., Warren E., Ward F.E., Parham P.;
RT   "Comparison of class I MHC alleles in humans and apes.";
RL   Immunol. Rev. 113:147-185(1990).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30678; AAA87970.1; -.
DR   PIR; I36961; I36961.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-2 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   365 AA;  40848 MW;  FC452786BD038D3E CRC64;
```

```
  Query Match             33.8%;  Score 511;  DB 1;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 1.6e-35;
  Matches  110;  Conservative   44;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
Db         85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 11
HA1B_BOVIN
ID   HA1B_BOVIN     STANDARD;      PRT;   364 AA.
AC   P13753;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   BOLA class I histocompatibility antigen, alpha chain BL3-7 precursor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258075; PubMed=3133413;
RA   Ennis P.D., Jackson A.P., Parham P.;
RT   "Molecular cloning of bovine class I MHC cDNA.";
RL   J. Immunol. 141:642-651(1988).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21043; AAA30641.1; -.
DR   HSSP; P16391; 1ED3.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28    364       BOLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
```

```
FT                                    ALPHA CHAIN BL3-7.
FT   DOMAIN       28    117           EXTRACELLULAR ALPHA-1.
FT   DOMAIN      118    209           EXTRACELLULAR ALPHA-2.
FT   DOMAIN      210    301           EXTRACELLULAR ALPHA-3.
FT   DOMAIN      302    310           CONNECTING PEPTIDE.
FT   TRANSMEM    311    331
FT   DOMAIN      332    364           CYTOPLASMIC.
FT   CARBOHYD    106    106           N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113           N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    128    191           BY SIMILARITY.
FT   DISULFID    230    286           BY SIMILARITY.
SQ   SEQUENCE   364 AA;  41513 MW;  622056CF7DCF7873 CRC64;

  Query Match             33.7%;  Score 510;  DB 1;  Length 364;
  Best Local Similarity   38.9%;  Pred. No. 1.9e-35;
  Matches  109;  Conservative   49;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISS 59
              |  |||| | : | |    ||    | |:|||||  || :| :  : | ||| ||:
Db         26 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 84

Qy         60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
              : | : ::  |    :| || |:  :| |:  || :| : ||::  |     |:|::|
Db         85 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 144

Qy        118 YDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
              |||:| :    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db        145 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 203

Qy        178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
                | :  ||   |||| ::    |||| || :||: |:: | :: :    :: |  :  |
Db        204 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 262

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              :||||:|  |   | || ||||||||:|:| || :|| : ||
Db        263 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 302


RESULT 12
1A11_HUMAN
ID   1A11_HUMAN     STANDARD;      PRT;   365 AA.
AC   P13746;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HLA class I histocompatibility antigen, A-11 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*1101/A*1102).
RX   MEDLINE=89030641; PubMed=2460344;
RA   Mayer W.E., Jonker M., Klein D., Ivanyi P., van Seventer G.,
RA   Klein J.;
RT   "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for
RT   trans-species mode of evolution.";
RL   EMBO J. 7:2765-2774(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (A*1101/A*1102).
RX   MEDLINE=94287401; PubMed=8016845;
RA   Lin L., Tokunaga K., Ishikawa Y., Bannai M., Kashiwase K.,
RA   Kuwata S., Akaza T., Tadokoro K., Shibata Y., Juji T.;
RT   "Sequence analysis of serological HLA-A11 split antigens, A11.1 and
RT   A11.2.";
RL   Tissue Antigens 43:78-82(1994).
RN   [3]
RP   SEQUENCE OF 26-365 FROM N.A. (A*1101).
RX   MEDLINE=87192928; PubMed=2437024;
RA   Cowan E.P., Jelachich M.L., Biddison W.E., Coligan J.E.;
RT   "DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows
RT   identification of residues involved in epitopes recognized by
RT   antibodies and T cells.";
RL   Immunogenetics 25:241-250(1987).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-11 ARE KNOWN: A*1101 (A-
CC       11E) AND A*1102 (A-11K). THE SEQUENCE SHOWN IS THAT OF A*1101.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13111; CAA31503.1; -.
DR   EMBL; X13112; CAA31504.1; -.
DR   EMBL; D16841; BAA04117.1; -.
DR   EMBL; D16842; BAA04118.1; -.
DR   EMBL; M16010; AAA65449.1; -.
DR   EMBL; M16007; AAA65449.1; JOINED.
DR   EMBL; M16008; AAA65449.1; JOINED.
DR   EMBL; M16009; AAA65449.1; JOINED.
DR   PIR; I83063; I83063.
DR   PIR; S03536; A47636.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365           HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                     A-11 ALPHA CHAIN.
FT   DOMAIN       25    114           EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206           EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298           EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308           CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365           CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110           N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188           BY SIMILARITY.
FT   DISULFID    227    283           BY SIMILARITY.
FT   VARIANT      43     43           E -> K (IN ALLELE A*1102).
FT                                     /FTId=VAR_004353.
SQ   SEQUENCE   365 AA;  40937 MW;  FE449CE2D4BF6CC5 CRC64;

  Query Match             33.6%;  Score 508;  DB 1;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 2.8e-35;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
```

```
            :| :    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 13
1A03_HUMAN
ID   1A03_HUMAN     STANDARD;      PRT;   370 AA.
AC   P04439;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HLA class I histocompatibility antigen, A-3 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*0301).
RX   MEDLINE=84207948; PubMed=6609814;
RA   Strachan T., Sodoyer R., Damotte M., Jordan B.R.;
RT   "Complete nucleotide sequence of a functional class I HLA gene,
RT   HLA-A3: implications for the evolution of HLA genes.";
RL   EMBO J. 3:887-894(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (A*0302).
RX   MEDLINE=85290871; PubMed=2993417;
RA   Cowan E.P., Jordan B.E., Coligan J.E.;
RT   "Molecular cloning and DNA sequence analysis of genes encoding
RT   cytotoxic T lymphocyte-defined HLA-A3 subtypes: the E1 subtype.";
RL   J. Immunol. 135:2835-2841(1985).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-3 ARE KNOWN: A*0301 (A-
CC       3.1) AND A*0302. THE SEQUENCE SHOWN IS THAT OF A*0301.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00492; CAA25162.1; ALT_TERM.
DR   PIR; A02192; HLHUA3.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     29
FT   CHAIN         30    370       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-3 ALPHA CHAIN.
FT   DOMAIN        30    119       EXTRACELLULAR ALPHA-1.
FT   DOMAIN        120   211       EXTRACELLULAR ALPHA-2.
FT   DOMAIN        212   303       EXTRACELLULAR ALPHA-3.
FT   DOMAIN        304   313       CONNECTING PEPTIDE.
FT   TRANSMEM      314   337
FT   DOMAIN        338   370       CYTOPLASMIC TAIL.
FT   CARBOHYD      115   115       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID      130   193       BY SIMILARITY.
FT   DISULFID      232   288       BY SIMILARITY.
FT   VARIANT       181   181       E -> V (IN ALLELE A*0302).
FT                                 /FTId=VAR_004351.
FT   VARIANT       185   185       L -> Q (IN ALLELE A*0302).
FT                                 /FTId=VAR_004352.
SQ   SEQUENCE   370 AA;  41368 MW;  ABB1FA77460318A2 CRC64;

  Query Match             33.4%;  Score 506;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 4.2e-35;
  Matches  110;  Conservative   46;  Mismatches  112;  Indels   10;  Gaps    8;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |   |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db         90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              :| :   |  | ||:  |  || :||  |:  | | |||:  |   |:: || |:
Db        150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy        180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db        208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||:| |  : || |||||||| |:| || :|| : ||
Db        267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 14
1A80_HUMAN
ID   1A80_HUMAN     STANDARD;      PRT;   365 AA.
AC   Q09160;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, AW-80(A-1) alpha chain
DE   precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94245293; PubMed=8188325;
RA   Balas A., Garcia-Sanchez F., Gomez-Reino F., Vicario J.L.;
RT   "Characterization of a new and highly distinguishable HLA-A allele in
RT   a Spanish family.";
RL   Immunogenetics 39:452-452(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Domena J.D.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE ONLY ALLELE OF AW-80 KNOWN IS A*8001 WHICH IS
```

```
CC       SHOWN HERE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U03754; AAC04322.1; -.
DR   EMBL; L18898; AAA17012.1; -.
DR   PIR; I59638; I38439.
DR   HSSP; Q95352; 1HHK.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 AW-80(A-1) ALPHA CHAIN.
FT   DOMAIN        25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN       115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    332
FT   DOMAIN       333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD     110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID     125    188       BY SIMILARITY.
FT   DISULFID     227    283       BY SIMILARITY.
SQ   SEQUENCE   365 AA;  40791 MW;  CE1BC1CD60CA8FA8 CRC64;

  Query Match           33.3%;  Score 504;  DB 1;  Length 365;
  Best Local Similarity  38.3%;  Pred. No. 6.1e-35;
  Matches  106;  Conservative   52;  Mismatches  111;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :|: :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::::|        :  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||  : || | ||||| :|   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| | : || |:|:|||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 15
1A31_HUMAN
ID   1A31_HUMAN     STANDARD;      PRT;   365 AA.
AC   P16189; O98137; Q9TQ24;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HLA class I histocompatibility antigen, A-31 alpha chain precursor
DE   (Aw-19).
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*3101).
RX   MEDLINE=90038496; PubMed=2478623;
RA   Kato K., Trapani J.A., Allopenna J., Dupont B., Yang S.Y.;
RT   "Molecular analysis of the serologically defined HLA-Aw19 antigens. A
RT   genetically distinct family of HLA-A antigens comprising A29, A31,
RT   A32, and Aw33, but probably not A30.";
RL   J. Immunol. 143:3371-3378(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (A*3101).
RX   MEDLINE=92269955; PubMed=1317015;
RA   Belich M.P., Madrigal J.A., Hildebrand W.H., Zemmour J.,
RA   Williams R.C., Luz R., Petzl-Erler M.L., Parham P.;
RT   "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL   Nature 357:326-329(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (A*31012).
RX   MEDLINE=96387675; PubMed=8795145;
RA   Arnett K.L., Adams E.J., Parham P.;
RT   "On the sequence of A*3101.";
RL   Tissue Antigens 47:428-430(1996).
RN   [4]
RP   SEQUENCE OF 9-365 FROM N.A. (A*3101).
RX   MEDLINE=92269956; PubMed=1589035;
RA   Watkins D.I., McAdam S.N., Liu X., Stang C.R., Milford E.L.,
RA   Levine C.G., Garber T.L., Dogon A.L., Lord C.I., Ghim S.H.,
RA   Troup G.M., Hughes A.L., Letvin N.L.;
RT   "New recombinant HLA-B alleles in a tribe of South American
RT   Amerindians indicate rapid evolution of MHC class I loci.";
RL   Nature 357:329-333(1992).
RN   [5]
RP   SEQUENCE FROM N.A. (A*3104).
RA   Bettinotti M.P., Dhillon G., Hackett J., Simonis T.B., Marincola F.M.;
RT   "A New HLA-A*31 allele.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 26-206 FROM N.A. (A*3104).
RA   Mitsuishi Y.;
RT   "New HLA-A31 allele identified in African American population.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-31 ARE KNOWN: A*3101 AND
CC       A*3104. THE SEQUENCE SHOWN IS THAT OF A*3101.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30578; AAA59613.1; -.
DR   EMBL; M84375; AAA59599.1; -.
DR   EMBL; L78918; AAB05976.1; -.
DR   EMBL; AF148863; AAD39981.1; -.
DR   EMBL; AF105028; AAC79721.1; -.
DR   EMBL; AF105027; AAC79721.1; JOINED.
DR   PIR; I72170; I72170.
DR   HSSP; O19673; 1HSB.
```

```
DR   MIM; 142800; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-31 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   VARIANT     121    121       M -> I (IN ALLELE A*3104).
FT                                 /FTId=VAR_010373.
FT   VARIANT     138    138       Q -> R (IN ALLELE A*3104).
FT                                 /FTId=VAR_010374.
SQ   SEQUENCE   365 AA;  41004 MW;  4E760C821A3C553B CRC64;

  Query Match             33.2%;  Score 502;  DB 1;  Length 365;
  Best Local Similarity   39.0%;  Pred. No. 8.9e-35;
  Matches  108;  Conservative   50;  Mismatches  111;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: |      |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFTTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|    :  ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db         85 DQETRNVKAHSQIDRVDLGTLRGYYNQSEAGSHTIQMMYGCDVGSDGRFLRGYQQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  |: :||  :: | | ||||| |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITQRKWEAARV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| | :: || |:|||||| |:| || :|| : ||
Db        263 GTFQKWASVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: August  5, 2003, 13:08:49
Job time : 10.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:05:29 ; Search time 32 Seconds
(without alignments)
2225.704 Million cell updates/sec

Title: US-10-092-404-3
Perfect score: 1514
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1129	74.6	358	11	Q8C2A6	Q8c2a6 mus musculu
2	1129	74.6	359	11	Q9D754	Q9d754 mus musculu
3	792	52.3	272	11	Q9R105	Q9r105 rattus norv
4	574	37.9	116	4	Q9HC69	Q9hc69 homo sapien
5	540.5	35.7	359	7	Q8HX81	Q8hx81 ornithorhyn
6	537.5	35.5	354	7	Q95HB3	Q95hb3 anas platyr
7	531.5	35.1	340	7	Q9BD50	Q9bd50 pongo pygma
8	530.5	35.0	334	7	Q9TQK3	Q9tqk3 homo sapien
9	530.5	35.0	341	4	Q9NPL2	Q9npl2 homo sapien
10	530.5	35.0	341	7	Q95460	Q95460 homo sapien
11	530.5	35.0	341	7	Q9BCU3	Q9bcu3 pan troglod
12	527.5	34.8	341	7	Q9BCU4	Q9bcu4 pan troglod
13	515	34.0	356	7	Q8HX66	Q8hx66 sus scrofa
14	514	33.9	332	7	Q30990	Q30990 pan troglod
15	514	33.9	365	7	Q9TPL7	Q9tpl7 pan troglod
16	512	33.8	105	4	Q9HC71	Q9hc71 homo sapien
17	512	33.8	129	4	Q9UK37	Q9uk37 homo sapien
18	512	33.8	365	7	Q9TQP6	Q9tqp6 homo sapien
19	512	33.8	371	7	Q9TQP7	Q9tqp7 homo sapien
20	511	33.8	352	7	Q8MHT1	Q8mht1 sus scrofa
21	511	33.8	364	7	O19243	O19243 sus scrofa
22	511	33.8	365	7	Q9MYG4	Q9myg4 homo sapien
23	509	33.6	365	7	Q29747	Q29747 homo sapien
24	508	33.6	273	7	Q95IG6	Q95ig6 homo sapien
25	508	33.6	352	7	Q8SPA9	Q8spa9 sus scrofa
26	508	33.6	361	7	Q8HX63	Q8hx63 sus scrofa
27	508	33.6	364	7	Q8HX61	Q8hx61 sus scrofa
28	508	33.6	365	7	Q9MYI5	Q9myi5 homo sapien
29	507	33.5	91	4	Q9BXI5	Q9bxi5 homo sapien
30	507	33.5	91	7	Q8MH46	Q8mh46 homo sapien
31	507	33.5	330	7	O19356	O19356 macaca mula
32	507	33.5	331	7	O02944	O02944 macaca mula
33	507	33.5	333	7	O98030	O98030 papio anubi
34	507	33.5	333	7	O98031	O98031 papio anubi
35	506	33.4	330	7	O02947	O02947 macaca mula
36	506	33.4	330	7	O02946	O02946 macaca mula
37	506	33.4	365	7	O19756	O19756 homo sapien
38	505	33.4	331	7	O02945	O02945 macaca mula
39	505	33.4	357	7	Q30886	Q30886 pan paniscu
40	505	33.4	363	7	Q9MXI5	Q9mxi5 pan troglod
41	505	33.4	363	7	Q9MWK4	Q9mwk4 gorilla gor
42	505	33.4	365	7	Q9MXI6	Q9mxi6 pan troglod
43	505	33.4	365	7	Q9MXM7	Q9mxm7 pan troglod
44	504	33.3	360	7	Q9GJ24	Q9gj24 homo sapien
45	504	33.3	364	7	Q8SPA4	Q8spa4 sus scrofa

ALIGNMENTS

RESULT 1
Q8C2A6
ID Q8C2A6 PRELIMINARY; PRT; 358 AA.
AC Q8C2A6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hemochromatosis.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";.
RL Nature 420:563-573(2002).
DR EMBL; AK088986; BAC40688.1; -.
SQ SEQUENCE 358 AA; 40421 MW; EE88FB6E5AAC844D CRC64;

Query Match 74.6%; Score 1129; DB 11; Length 358;
Best Local Similarity 71.9%; Pred. No. 1.3e-97;
Matches 202; Conservative 30; Mismatches 41; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |:||||| ||| ||:  : |||:||
Db         29 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 88

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              |||||||||:|| ||||||| |:| ||         ||| |||:|||| ||||| |:|:
Db         89 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 148

Qy        116 YGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              ||||||| ||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        149 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 208
```

```
Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::||||||:|||| ||:|||:  |: ||
Db        209 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 268

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:|||||||||||||   ||
Db        269 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 309
```

RESULT 2
Q9D754
ID Q9D754 PRELIMINARY; PRT; 359 AA.
AC Q9D754;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310032M04, full insert sequence.
GN HFE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AK009581; BAB26373.1; -.
DR HSSP; Q30201; 1A6Z.
DR MGD; MGI:109191; Hfe.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 359 AA; 40534 MW; 586657B7F9FF20B4 CRC64;

Query Match 74.6%; Score 1129; DB 11; Length 359;
Best Local Similarity 71.9%; Pred. No. 1.3e-97;
Matches 202; Conservative 30; Mismatches 41; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| |||||||||  |:||||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              |||||||||:|| |||||| |:| ||         ||| |||:|||| ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              ||||||| ||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::||||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:|||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310
```

RESULT 3
Q9R105
ID Q9R105 PRELIMINARY; PRT; 272 AA.
AC Q9R105;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hemochromatosis gene product HFE splice variant delE2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Testis;
RA Liew Y.-F., Shaw N.-S.;
RT "Alternative splice variant of the hemochromatosis gene HFE in iron
RT overloaded rats.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF176534; AAD49965.1; -.
DR HSSP; Q30201; 1A6Z.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 272 AA; 30757 MW; 1D91063CCBEF5502 CRC64;

Query Match 52.3%; Score 792; DB 11; Length 272;
Best Local Similarity 74.6%; Pred. No. 3.7e-66;
Matches 138; Conservative 22; Mismatches 25; Indels 0; Gaps 0;

```
Qy         92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHK 151
              ||| |||||||: ||||| |:|||||||| ||||| ||:| ||||||| ||:||| |:
Db         39 ESHILQVILGCEVHEDNSTSGFWKYGYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHR 98

Qy        152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQ 211
              |||||:| ||:|||| ||:|:||| |||| |||| |||||  |  |: |:|||:|||::||
Db         99 IRARQSRDYLQRDCPQQLKQVLELQRGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQ 158

Qy        212 NITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPL 271
```

```
                 ||||:|||| ||:|||:  |::|||||||||||||:|||| |||| |::||||||||||||
Db        159 NITMRWLKDSQPLDAKDVNPENVLPNGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPL 218

Qy        272 IVIWE 276
                 ||
Db        219 TATWE 223


RESULT 4
Q9HC69
ID   Q9HC69      PRELIMINARY;      PRT;   116 AA.
AC   Q9HC69;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hemochromatosis splice variant 861-2305del (Fragment).
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
DR   EMBL; AF144241; AAG29576.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   116 AA;  13541 MW;  AC0333B096A3F47B CRC64;

  Query Match             37.9%;  Score 574;  DB 4;  Length 116;
  Best Local Similarity   98.1%;  Pred. No. 3.6e-46;
  Matches  104;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         90 SKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWER 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 SKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWER 61

Qy        150 HKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVT 195
              |||||||||||||||||||||||||||||||||||||||||||||:
Db         62 HKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVS 107


RESULT 5
Q8HX81
ID   Q8HX81      PRELIMINARY;      PRT;   359 AA.
AC   Q8HX81;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I antigen.
OS   Ornithorhynchus anatinus (Duckbill platypus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX   NCBI_TaxID=9258;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22226695; PubMed=12242589;
RA   Miska K.B., Harrison G.A., Hellman L., Miller R.D.;
RT   "The major histocompatibility complex in monotremes: an analysis of
RT   the evolution of Mhc class I genes across all three mammalian
RT   subclasses.";
RL   Immunogenetics 54:381-393(2002).
DR   EMBL; AY112715; AAM54213.1; -.
SQ   SEQUENCE   359 AA;  39847 MW;  3FC701E44FD53B66 CRC64;

  Query Match             35.7%;  Score 540.5;  DB 7;  Length 359;
  Best Local Similarity   40.6%;  Pred. No. 2.1e-42;
  Matches  113;  Conservative   51;  Mismatches  107;  Indels    7;  Gaps    5;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHES--RRVEPRTPWVSSRISSQM 61
              ||||| | : | |    |:| | |:||||:  || :: :    | ||||||:      :
Db         25 RSHSLRYFYTGVSRPGSGVSEFTAVGYVDEVQFVRFEGDGGRGRAEPRTPWIRDNEGQEY 84

Qy         62 WLQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNST-EGYWKYGYD 119
              | : :| ||    :|  :  | :  :| ||| ||| | : ||:::||::  :|: :||||
Db         85 WDRNTQILKDTAQVFRTNLQTALGYYNQSKEGSHTFQRMYGCQLREDDTPGQGFTQYGYD 144

Qy        120 GQDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              |:| :     || | ||:  |  || :||  :  | :::|||| :|   |:: :  |
Db        145 GRDYIALDKATLSWTAADAGALNTKRKWEADRTIAERDKAYLEDECIEWLRKYVGFGGPS 204

Qy        180 LDQQVPPLVKVTHHV--TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPN 237
              | :  || |:|| |         :|||||| :|| :|  || :: : | ::| |     |:
Db        205 LTRADPPEVEVTRHTGPDGEDVSLRCRALGFYPADIKFKWEREGKDM-SQEMEFVGTRPS 263

Qy        238 GDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
              ||| :| | :: || ||||:| | ||| || ||| | |
Db        264 GDGNFQKWASVNVPRGEEQKYVCVVEHEGLGQPLAVKW 301


RESULT 6
Q95HB3
ID   Q95HB3      PRELIMINARY;      PRT;   354 AA.
AC   Q95HB3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I antigen alpha chain (Fragment).
OS   Anas platyrhynchos (Domestic duck).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX   NCBI_TaxID=8839;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF393511; AAK84356.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; MHC; Transmembrane.
FT   NON_TER       1       1
SQ   SEQUENCE   354 AA;  39405 MW;  D44FACEE7B60468F CRC64;

  Query Match             35.5%;  Score 537.5;  DB 7;  Length 354;
  Best Local Similarity   37.3%;  Pred. No. 3.9e-42;
  Matches  101;  Conservative   60;  Mismatches  109;  Indels    1;  Gaps    1;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQL 65
              ||| |     |:   |:  | |:||| ::|| || |:| :|||  |::  :  | |
Db         24 HSLRYFATAVSDPSPGVPQFVAVGYVDGEVFVRYDSETRGMEPRVDWIADNMDQQYWNGE 83

Qy         66 SQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALE 125
              :::|:| : :: ||  |: | :| |: ||||| : ||:: || |  |:::|||:|:: :
Db         84 TENLRGAEQIYRVDLETLRERYNQSRGSHTLQHMFGCDLLEDRSISGFFQYGYEGREFIA 143
```

```
Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              ||  : ||:  |  || :||    | : : |||  |   |:: :  |: ||:::
Db        144 LDKDTWTFTAADAAAQITKRKWEEDGTVAERRKYYLENTCIEWLRKYVRYGKDVLERRER 203

Qy        186 PLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGW 245
              | |:|:      : :| |||  :||: |:: ||||   :  :| :    :|| ||||  |
Db        204 PEVRVSGMEADKILSLSCRAHGFYPRPISISWLKDGM-VQEQETQRGSTVPNSDGTYHIW 262

Qy        246 ITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |: | ||:: :| |:|||  | || :  ||
Db        263 ATIDVVPGDKDKYQCRVEHASLPQPGLFSWE 293


RESULT 7
Q9BD50
ID   Q9BD50      PRELIMINARY;      PRT;   340 AA.
AC   Q9BD50;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I related protein MR1 isoform.
GN   MR1.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.,
RA   Martinez-Naves E.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ271828; CAC28215.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   340 AA;  39375 MW;  A893952B78725F17 CRC64;

  Query Match             35.1%;  Score 531.5;  DB 7;  Length 340;
  Best Local Similarity   39.1%;  Pred. No. 1.4e-41;
  Matches  106;  Conservative   51;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||       ||  ::: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIRGVPEFISVGYVDSHPITTYDSVTQQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::   :  : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKRAWEANQHELQYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   |||| |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTTLFCKAHGFYPPEIYMTWMKNGEEI-VQEMDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :  : |     |:| ||| |:   | |
Db        261 QTWASFELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 8
Q9TQK3
ID   Q9TQK3      PRELIMINARY;      PRT;   334 AA.
AC   Q9TQK3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I-related protein MR1 (Fragment).
GN   MR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=99003494; PubMed=9784382;
RA   Yamaguchi H., Kurosawa Y., Hashimoto K.;
RT   "Expanded genomic organization of conserved mammalian MHC class I-
RT   related genes, human MR1 and its murine ortholog.";
RL   Biochem. Biophys. Res. Commun. 250:558-564(1998).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF073485; AAC72900.1; -.
DR   EMBL; AF073484; AAC72900.1; JOINED.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   334 AA;  38586 MW;  4C3E3A8248A39BA4 CRC64;

  Query Match             35.0%;  Score 530.5;  DB 7;  Length 334;
  Best Local Similarity   39.1%;  Pred. No. 1.6e-41;
  Matches  106;  Conservative   50;  Mismatches  112;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||       ||  :|: ||| ||::  ::   |
Db         16 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 75

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         76 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 134

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db        135 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 194

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   ||| |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        195 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 253

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        254 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 284


RESULT 9
Q9NPL2
```

```
ID   Q9NPL2      PRELIMINARY;      PRT;   341 AA.
AC   Q9NPL2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mr1 protein.
GN   MR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood;
RX   MEDLINE=20470599; PubMed=11019920;
RA   Parra-Cuadrado J.F., Navarro P. ., Mirones I., Setien F., Oteo M.,
RA   Martinez-Naves E.;
RT   "A study on the polymorphism of human MHC class I-related MR1 gene and
RT   identification of an MR1-like pseudogene.";
RL   Tissue Antigens 56:170-172(2000).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ249778; CAB77667.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   341 AA;  39366 MW;  2990C1F3F0A1CAD9 CRC64;

  Query Match             35.0%;  Score 530.5;  DB 4;  Length 341;
  Best Local Similarity   39.1%;  Pred. No. 1.7e-41;
  Matches  106;  Conservative   50;  Mismatches  112;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||       ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|:. : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 10
Q95460
ID   Q95460      PRELIMINARY;      PRT;   341 AA.
AC   Q95460;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Class I histocompatibility antigen-like protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=95350662; PubMed=7624800;
RA   Hashimoto K., Hirai M., Kurosawa Y.;
RT   "A gene outside the human MHC related to classical HLA class I
RT   genes.";
RL   Science 269:693-695(1995).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; U22963; AAC50174.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   Genew; HGNC:4975; HLALS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   341 AA;  39366 MW;  2990C1F3F0A1CAD9 CRC64;

  Query Match             35.0%;  Score 530.5;  DB 7;  Length 341;
  Best Local Similarity   39.1%;  Pred. No. 1.7e-41;
  Matches  106;  Conservative   50;  Mismatches  112;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||       ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|:. : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 11
Q9BCU3
ID   Q9BCU3      PRELIMINARY;      PRT;   341 AA.
AC   Q9BCU3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I related protein, MR1B1 isoform.
GN   MR1.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martinez-Naves E.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
```

```
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ275984; CAC34272.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   VARIANT     197    197       I -> T (IN REF. 2).
SQ   SEQUENCE   341 AA;  39394 MW;  FBFA822BCAB2C7A8 CRC64;

  Query Match            35.0%;  Score 530.5;  DB 7;  Length 341;
  Best Local Similarity  39.1%;  Pred. No. 1.7e-41;
  Matches  106;  Conservative   51;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |: :| :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDILQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QTWASVELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 12
Q9BCU4
ID   Q9BCU4      PRELIMINARY;      PRT;   341 AA.
AC   Q9BCU4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I related protein, MR1B1 isoform.
GN   MR1.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martinez-Naves E.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ275982; CAC34274.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   VARIANT     197    197       T -> I (IN REF. 2).
SQ   SEQUENCE   341 AA;  39382 MW;  DFF16AF1FAB2D272 CRC64;

  Query Match            34.8%;  Score 527.5;  DB 7;  Length 341;
  Best Local Similarity  39.1%;  Pred. No. 3.2e-41;
  Matches  106;  Conservative   50;  Mismatches  112;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDA 123
              : :| |:||  || |:   :  ::| |  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QTWASVELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 13
Q8HX66
ID   Q8HX66      PRELIMINARY;      PRT;   356 AA.
AC   Q8HX66;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I antigen (Fragment).
GN   SLA-1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martens G.W., Baker J.E., Smith D.M.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY135589; AAN35107.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   356 AA;  39585 MW;  94FC7A461DBF555B CRC64;

  Query Match            34.0%;  Score 515;  DB 7;  Length 356;
  Best Local Similarity  39.9%;  Pred. No. 5.1e-40;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMWL 63
              ||| | :   |  ||| | | |:|||||  || :| :  : |:||| ||:       : |
Db         19 HSLRYFYTAVSRPDLGDSRFIAVGYVDDTQFVRFDSDAPNPRMEPRAPWIQQE-GQEYWD 77

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDN-STEGYWKYGYDGQ 121
```

```
Db         78 EETRNAMGSAQNDRVDLKTLRGYYNQSEAGSHTIQRMYGCDVGPDGLLLRGYDQDAYDGA 137

Qy        122 DALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181

Db        138 DYIALNEDLRSWTAADTAAQITKRKWEAANV-AEQERSYLEGTCVEWLQKYLEMGKDTLQ 196

Qy        182 QQVPPLVKVTHHVTSSV-TTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Db        197 RAEPPKTHVTRHPSSDLGVTLRCWALGFYPKEISLTWQREGQD-QSQDMELVETRPSGDG 255

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276

Db        256 TFQKWAALVVPPGEEQSYTCHVQHEGLQEPLTLRWD 291


RESULT 14
Q30990
ID   Q30990      PRELIMINARY;      PRT;   332 AA.
AC   Q30990;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chimpanzee MHC class I ChlA chain (Fragment).
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89235215; PubMed=2715640;
RA   Parham P., Lawlor D.A., Lomen C.E., Ennis P.D.;
RT   "Diversity and diversification of HLA-A,B,C alleles.";
RL   J. Immunol. 142:3937-3950(1989).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; M24047; AAA35426.1; -.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37433 MW;  9AA9A55DF9E79360 CRC64;

  Query Match             33.9%;  Score 514;  DB 7;  Length 332;
  Best Local Similarity   40.1%;  Pred. No. 5.8e-40;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62

Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120

Db         85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239

Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276

Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
Q9TPL7
ID   Q9TPL7      PRELIMINARY;      PRT;   365 AA.
AC   Q9TPL7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   MHC class I antigen.
GN   HLA-A OR PATR-A.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=35A-1;
RX   MEDLINE=99335357; PubMed=10405321;
RA   Matsui M., Machida S., Feinstone S.M., Akatsuka T.;
RT   "Molecular analyses of five new chimpanzee MHC class I alleles:
RT   implications for differences between evolutional mechanisms of HLA-A,
RT   -B, and -C loci.";
RL   Biochem. Biophys. Res. Commun. 261:46-52(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=20322475; PubMed=10866106;
RA   de Groot N.G., Otting N., Arguello R., Watkins D.I., Doxiadis G.G.M.,
RA   Madrigal J.A., Bontrop R.E.;
RT   "Major histocompatibility complex class I diversity in a West African
RT   chimpanzee population: implications for HIV research.";
RL   Immunogenetics 51:398-409(2000).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF115459; AAF02438.1; -.
DR   EMBL; AF168393; AAF72774.1; -.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; MHC; Transmembrane.
SQ   SEQUENCE   365 AA;  40819 MW;  4E85F08E33479E38 CRC64;

  Query Match             33.9%;  Score 514;  DB 7;  Length 365;
  Best Local Similarity   40.1%;  Pred. No. 6.5e-40;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62

Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120

Db         85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
```

```
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299

Search completed: August  5, 2003, 13:10:01
Job time : 33 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:05:29 ; Search time 38 Seconds
(without alignments)
1152.856 Million cell updates/sec

Title: US-10-092-404-3
Perfect score: 1514
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1514	100.0	276	20	AAW94297	HFE mutant (H111A/
2	1514	100.0	276	24	ABG72687	Human haemochromat
3	1502	99.2	276	20	AAW94295	Wild-type HFE poly
4	1502	99.2	276	24	ABG72685	Human haemochromat
5	1502	99.2	348	18	AAW36499	Hereditary haemoch
6	1502	99.2	348	21	AAB19149	A human histocompa
7	1502	99.2	348	22	AAB36869	Human hereditary h
8	1497	98.9	438	23	AAU80035	Beta 2 microglobul
9	1493	98.6	276	20	AAW94296	HFE mutant (H63D-H
10	1493	98.6	276	24	ABG72686	Human haemochromat
11	1493	98.6	348	22	AAB36871	Human hereditary h
12	1491	98.5	348	22	AAB36870	Human hereditary h
13	1482	97.9	348	22	AAB36872	Human hereditary h
14	517	34.1	361	22	AAB36873	Rabbit leukocyte a
15	508	33.6	365	22	AAB36874	MHC class I protei
16	504	33.3	92	24	ABP68379	Human colon specif
17	500	33.0	274	21	AAY68275	Human leukocyte an
18	500	33.0	274	21	AAY52929	HLA-A2/A28 family
19	500	33.0	274	22	AAB58690	HLA-A2/A28 protein
20	500	33.0	280	22	AAU10225	Human leukocyte an
21	500	33.0	280	24	ABU08672	Human histocompati
22	500	33.0	415	22	AAU10224	Human partial beta
23	500	33.0	415	24	ABU08671	Human single chain
24	499	33.0	365	21	AAY68265	Human leukocyte an
25	499	33.0	365	21	AAY52919	HLA-A2/A28 family
26	499	33.0	365	22	AAB58680	HLA-A2/A28 protein
27	499	33.0	368	22	AAM24017	Human EST encoded
28	498	32.9	274	21	AAY68276	Human leukocyte an
29	498	32.9	274	21	AAY52930	HLA-A2/A28 family
30	498	32.9	274	22	AAB58691	HLA-A2/A28 protein
31	498	32.9	365	21	AAY68268	Human leukocyte an
32	498	32.9	365	21	AAY52922	HLA-A2/A28 family
33	498	32.9	365	22	AAB58683	HLA-A2/A28 protein
34	497	32.8	274	9	AAP80911	Consensus sequence
35	497	32.8	365	21	AAY68267	Human leukocyte an
36	497	32.8	365	21	AAY52921	HLA-A2/A28 family
37	497	32.8	365	22	AAB58682	HLA-A2/A28 protein
38	496	32.8	274	21	AAY68274	Human leukocyte an
39	496	32.8	274	21	AAY52928	HLA-A2/A28 family
40	496	32.8	274	22	AAB58689	HLA-A2/A28 protein
41	496	32.8	365	21	AAY68266	Human leukocyte an
42	496	32.8	365	21	AAY52920	HLA-A2/A28 family
43	496	32.8	365	22	AAB58681	HLA-A2/A28 protein
44	495	32.7	412	19	AAW68385	Chimeric HLA-A2.1/
45	494	32.6	274	21	AAY68273	Human leukocyte an

ALIGNMENTS

RESULT 1
AAW94297
ID AAW94297 standard; peptide; 276 AA.
XX
AC AAW94297;
XX
DT 27-APR-1999 (first entry)
XX
DE HFE mutant (H111A/H145A-HFE) polypeptide sequence.
XX
KW HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW transfusion; protein replacement therapy; hereditary hemochromatosis;
KW transferrin receptor; iron deficiency; anemia; mutant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "indicated in the sequence listing as Arg"
FT Misc-difference 89
FT /label= H111A
FT /note= "wild type His (of the mature protein sequence)
FT is replaced by Ala"
FT Misc-difference 123
FT /label= H145A
FT /note= "wild type His (of the mature protein sequence)
FT is replaced by Ala"
XX
PN WO9856814-A1.
XX
PD 17-DEC-1998.
XX

PF 12-JUN-1998; 98WO-US12436.
XX
PR 13-JUN-1997; 97US-0876010.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (PROG-) PROGENITOR INC.
XX
PI Bjorkman PJ, Feder JN, Schatzman RC;
XX
DR WPI; 1999-080886/07.
XX
PT New treatment of an iron overload disease - comprises use of HFE
PT polypeptides provided in a complex with full length, wild type human
PT (2m), useful in protein replacement therapy
XX
PS Claim 5; Page 15; 36pp; English.
XX
CC The present sequence represents a H111A/H145A-HFE mutant polypeptide.
CC The HFE polypeptides (AAW94295-297) provided in a complex with full
CC length, wild type human beta-2-microglobulin (beta2m) form compositions
CC in the treatment of primary iron overload diseases (e.g.
CC hemochromatosis), or other iron overload conditions resulting from
CC secondary causes (e.g. repeated transfusions). Data regarding the
CC structure and function correlations of HFE polypeptides is useful in
CC designing drugs that modulate the HFE gene and HFE activity. The
CC polypeptides are also useful in protein replacement therapy for
CC individuals possessing a defective HFE gene (e.g. Hereditary
CC hemochromatosis). (Ant)agonists of the polypeptides are also useful in
CC treating primary and secondary iron overload diseases. The modulators of
CC the transferrin receptor are useful in treating iron deficiency
CC conditions such as anemia, and in modulating the amount of iron
CC transported into a cell. The HFE polypeptides provide a molecular basis
CC for the relationship between HFE and iron metabolism, which enables
CC treatment of iron overload and deficiency diseases.
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1514; DB 20; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.1e-132;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
Db 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
Db 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy 121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
Db 121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
Db 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy 241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db 241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276

RESULT 2
ABG72687
ID ABG72687 standard; protein; 276 AA.
XX
AC ABG72687;
XX
DT 05-MAR-2003 (first entry)
XX
DE Human haemochromatosis (HFE) mature protein, mutant H89A/H123A.
XX
KW Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW Beta2m; transferrin receptor; anaemia; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 89
FT /note= "Wild-type His substituted by Ala"
FT Misc-difference 123
FT /note= "Wild-type His substituted by Ala"
XX
PN US6391852-B1.
XX
PD 21-MAY-2002.
XX
PF 12-JUN-1998; 98US-0094964.
XX
PR 13-JUN-1997; 97US-0876010.
XX
PA (BIRA) BIO-RAD LAB INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 2003-155377/15.
XX
PT Method of treating an iron overload disease comprises administration of
PT a soluble complex comprising a 276 amino acid HFE polypeptide and a
PT full length, wild-type human beta2m -
XX
PS Claim 3; Column 2; 17pp; English.
XX
CC The invention relates to a method of treating an iron overload disease
CC comprising administration of a soluble complex comprising a 276 amino
CC acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC (ABG72685-ABG72687) and a full length, wild-type human beta2m
CC (beta2-microglobulin). In a HeLa cell based assay, binding and uptake of
CC ^1^2^5I-transferrin in the presence of purified H63D-HFE/beta2m
CC heterodimers was determined. At a concentration of 250 nM H63D-HFE/
CC beta2m heterodimers, the transferrin receptor (TfR) displayed a KD for
CC transferrin of 28 nM. At the same concentration of normal HFE/beta_2m
CC heterodimers, TfR displayed a KD for transferrin of 40 nM. In the absence
CC of any HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of
CC 7nM. It was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC efficient in decreasing TfR affinity for transferrin compared to
CC wild-type HFE. The method is useful for treating iron overload diseases
CC and iron deficiency e.g. anaemia. The present sequence is the H111A/H145A
CC (residues 111 and 145 of the full length protein, 89/123 of the mature
CC form) mutant from of mature HFE used to investigate the role of the His
CC residues in transferrin receptor binding to transferrin.
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1514; DB 24; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.1e-132;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
Db 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
Db 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy 121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
Db 121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
Db 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

```
Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 3
AAW94295
ID   AAW94295 standard; peptide; 276 AA.
XX
AC   AAW94295;
XX
DT   27-APR-1999  (first entry)
XX
DE   Wild-type HFE polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US12436.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (PROG-) PROGENITOR INC.
XX
PI   Bjorkman PJ,  Feder JN,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
XX
PS   Claim 1; Page 13; 36pp; English.
XX
CC   The present sequence represents a wild-type HFE polypeptide. The HFE
CC   polypeptides (AAW94295-297)  provided in a complex with full length,
CC   wild type human beta-2-microglobulin (beta2m) form compositions in the
CC   treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC   other iron overload conditions resulting from secondary causes (e.g.
CC   repeated transfusions). Data regarding the structure and function
CC   correlations of HFE polypeptides is useful in designing drugs that
CC   modulate the HFE gene and HFE activity. The polypeptides are also useful
CC   in protein replacement therapy for individuals possessing a defective
CC   HFE gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the
CC   polypeptides are also useful in treating primary and secondary iron
CC   overload diseases. The modulators of the transferrin receptor are useful
CC   in treating iron deficiency conditions such as anemia, and in modulating
CC   the amount of iron transported into a cell. The HFE polypeptides provide
CC   a molecular basis for the relationship between HFE and iron metabolism,
CC   which enables treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match             99.2%;  Score 1502;  DB 20;  Length 276;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-131;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 4
ABG72685
ID   ABG72685 standard; protein; 276 AA.
XX
AC   ABG72685;
XX
DT   05-MAR-2003  (first entry)
XX
DE   Human haemochromatosis (HFE) mature protein.
XX
KW   Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW   iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW   Beta2m; transferrin receptor; anaemia.
XX
OS   Homo sapiens.
XX
PN   US6391852-B1.
XX
PD   21-MAY-2002.
XX
PF   12-JUN-1998;   98US-0094964.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (BIRA ) BIO-RAD LAB INC.
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI   Feder JN,  Bjorkman PJ,  Schatzman RC;
XX
DR   WPI; 2003-155377/15.
XX
PT   Method of treating an iron overload disease comprises administration of
PT   a soluble complex comprising a 276 amino acid HFE polypeptide and a
PT   full length, wild-type human beta2m -
XX
PS   Claim 1; Column 1; 17pp; English.
XX
CC   The invention relates to a method of treating an iron overload disease
CC   comprising administration of a soluble complex comprising a 276 amino
CC   acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC   (ABG72685-ABG72687) and a full length, wild-type human beta2m
CC   (beta2-microglobulin). In a HeLa cell based assay, binding and uptake of
CC   ^1^2^5I-transferrin in the presence of purified H63D-HFE/beta2m
CC   heterodimers was determined. At a concentration of 250 nM H63D-HFE/
CC   beta2m heterodimers, the transferrin receptor (TfR) displayed a KD for
CC   transferrin of 28 nM. At the same concentration of normal HFE/beta_2m
CC   heterodimers, TfR displayed a KD for transferrin of 40 nM. In the absence
CC   of any HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of
CC   7nM. It was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC   efficient in decreasing TfR affinity for transferrin compared to
CC   wild-type HFE. The method is useful for treating iron overload diseases
CC   and iron deficiency e.g. anaemia. The present sequence is wild-type
CC   mature HFE.
XX
SQ   Sequence   276 AA;
```

```
  Query Match             99.2%;  Score 1502;  DB 24;  Length 276;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-131;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 5
AAW36499
ID   AAW36499 standard; Protein; 348 AA.
XX
AC   AAW36499;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene product.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 63
FT                   /note= "substituted by Asp in 24s2 mutant"
FT   Misc-difference 65
FT                   /note= "substituted by Cys in 24d7 variant"
FT   Misc-difference 282
FT                   /note= "substituted by Tyr in 24d1 mutant"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   N-PSDB; AAT96690, AAT96691.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 4; 115pp; English.
XX
CC   This polypeptide is the expression product of a novel human gene
CC   (see AAT96690) whose mutated form is associated with hereditary
CC   haemochromatosis (HH).  A single mutation (24d1) in the HH gene
CC   appears responsible for the majority of HH disease.  This comprises
CC   a G to A substitution that is present in 86% of affected
CC   chromosomes and in 4% of unaffected chromosomes.  It results in a
CC   Cys to Tyr substitution in the encoded protein at a critical
CC   disulphide bridge important for secondary structure.  The following
CC   are claimed: the 10825 bp genomic DNA sequence (I), a 1437 bp cDNA
CC   sequence (Ia) (see AAT96691) and their 24d1, 24d2 and 24d7 variants;
CC   a cloning or expression vector; host cells; a peptide product
CC   chosen from the HH gene product, its variants (24d1, 24d2 and
CC   24d7), or a peptide of at least 56 amino acid residues of these; an
CC   antibody produced using the peptide as an immunogen; a method to
CC   determine the presence or absence of the common HH gene mutation;
CC   an animal model for the HH disease; metal chelation agents, T-cell
CC   differentiation factors and therapeutic agents for the mitigation
CC   of injury due to oxidative process in vivo or mitigation of iron
CC   overload; a method for screening potential therapeutic agents for
CC   activity in connection with HH disease; an antisense oligonucleotide
CC   directed against a transcriptional product of a nucleic acid
CC   sequence as above; and oligonucleotides or pairs of oligonucleotides
CC   covering a range of nucleotides from (I), (Ia) or their variants,
CC   useful for detecting a polymorphism in the HH gene.  The invention
CC   also relates to methods for screening for HH homozygotes, to HH
CC   diagnosis, prenatal screening and diagnosis, and therapies of HH
CC   disease, including gene therapy, protein- and antibody-based
CC   therapeutics, and small molecule therapeutics.
XX
SQ   Sequence   348 AA;

  Query Match             99.2%;  Score 1502;  DB 18;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 2e-131;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 6
AAB19149
ID   AAB19149 standard; Protein; 348 AA.
XX
AC   AAB19149;
XX
DT   19-FEB-2001  (first entry)
XX
DE   A human histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
```

```
FT                   /note= "signal peptide"
FT   Misc-difference 63
FT                   /note= "when nucleotide 187 is mutated to G, then
FT                          this residue is Asp"
FT   Misc-difference 65
FT                   /note= "when nucleotide 193 is mutated to T, then
FT                          this residue is Cys"
FT   Domain          80..108
FT                   /note= "alpha1 domain"
FT   Misc-difference 93
FT                   /note= "when nucleotide 277 is mutated to C, then
FT                          this residue is Arg"
FT   Misc-difference 105
FT                   /note= "when nucleotide 314 is mutated to C, then
FT                          this residue is Thr"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   N-PSDB; AAA96769.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Disclosure; Page 3; 55pp; English.
XX
CC   The present sequence represents a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence   348 AA;

  Query Match             99.2%;  Score 1502;  DB 21;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 2e-131;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 7
AAB36869
ID   AAB36869 standard; Protein; 348 AA.
XX
AC   AAB36869;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68425.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 1; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match             99.2%;  Score 1502;  DB 22;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 2e-131;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262
```

```
Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 8
AAU80035
ID   AAU80035 standard; Protein; 438 AA.
XX
AC   AAU80035;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Beta 2 microglobulin (beta2M)/HFE monochain.
XX
KW   Human; beta 2 microglobulin; beta2M/HFE monochain; HFE; ischaemia;
KW   iron absorption regulator; intracellular iron absorption; lung injury;
KW   haemochromatosis; transfusion; thalassaemia; haemolytic anaemia;
KW   chronic infection; transferrin receptor; TfR; brain tumour; cancer;
KW   oxidative stress disorder; tissue damage; vascular disease;
KW   inflammation; atherosclerosis; autoimmune disease;
KW   inflammatory condition.
XX
OS   Homo sapiens.
XX
PN   WO200224929-A2.
XX
PD   28-MAR-2002.
XX
PF   24-SEP-2001; 2001WO-US29873.
XX
PR   22-SEP-2000; 2000US-234843P.
XX
PA   (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
PA   (MCIN/) MCINNIS P.
XX
PI   Ehrlich R,  Rotem-Yehudar R,  Laham N;
XX
DR   WPI; 2002-383192/41.
DR   N-PSDB; ABK49917.
XX
PT   Soluble beta 2 microglobulin/HFE monochain useful for treating
PT   iron-overload conditions e.g. thalassemia and chronic infections,
PT   comprises human beta 2 microglobulin linked to alpha domains of HFE by
PT   a linker peptide  -
XX
PS   Example 2; Fig 2; 77pp; English.
XX
CC   The invention relates to a soluble polypeptide (I) of beta 2
CC   microglobulin (beta2m)/HFE monochain comprising human beta2m (or its
CC   analogue or active fragment), linked to alpha1-alpha3 domains of human
CC   HFE (a central regulator of iron absorption; undefined), or its analogue
CC   or active fragment, by a flexible linker peptide, or a functional
CC   derivative or salt of (I). (I) is useful for reducing intracellular iron
CC   absorption in patients having hereditary haemochromatosis, transfusions,
CC   thalassaemias, haemolytic anaemia or chronic infections, and for
CC   delivering a therapeutic to cells that over-express transferrin receptor
CC   (TfR) which are preferably lymphocytes or leukocytes, across the blood-
CC   brain barrier. (I) is further useful for treating brain tumour. (I)
CC   is also useful for treating oxidative stress disorders resulting in
CC   tissue damage e.g. vascular diseases, inflammation, atherosclerosis,
CC   lung injury, ischaemia, etc. A DNA molecule (II) encoding (I) is useful
CC   as a platform for drug delivery of therapeutic use for cancer,
CC   autoimmune diseases and inflammatory conditions. The monochain manifests
CC   specific characteristics advantageous for drug delivery systems. It is a
CC   soluble, stable and fully conformed protein. It binds specifically to
CC   transferrin receptor (TfR) and therefore targets cells that over-express
CC   this receptor. It is continuously internalised by the target cells, thus
CC   enabling efficient drug delivery. It dissociates from the receptor in the
CC   cells, minimising side effects. It negatively regulates iron absorption,
CC   reducing growth of undesired cells and preventing lymphocyte activation.
CC   It is not diluted in the blood as is transferrin. It should not induce an
CC   immune response since it is a self non-polymeric protein and delivery of
CC   drugs via monochain is expected to overcome drug-resistance since it is a
CC   natural TfR-binding protein. The present sequence represents the amino
CC   acid sequence of beta2m/HFE monochain.
XX
SQ   Sequence   438 AA;

  Query Match             98.9%;  Score 1497;  DB 23;  Length 438;
  Best Local Similarity   99.3%;  Pred. No. 7.8e-131;
  Matches  273;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 194

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        195 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 254

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 314

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 374

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
              |||||||||||||||||||||||||||||||||||
Db        375 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 409


RESULT 9
AAW94296
ID   AAW94296 standard; peptide; 276 AA.
XX
AC   AAW94296;
XX
DT   27-APR-1999  (first entry)
XX
DE   HFE mutant (H63D-HFE) polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia; mutant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
FT   Misc-difference 41
FT                   /label= H63D
FT                   /note= "wild type His (of the mature protein sequence)
FT                        is replaced by Asp"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US12436.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (PROG-) PROGENITOR INC.
XX
PI   Bjorkman PJ,  Feder JN,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
```

```
XX
PS   Claim 3; Page 14; 36pp; English.
XX
CC   The present sequence represents a H63D-HFE mutant polypeptide. The HFE
CC   polypeptides (AAW94295-297)  provided in a complex with full length,
CC   wild type human beta-2-microglobulin (beta2m) form compositions in the
CC   treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC   other iron overload conditions resulting from secondary causes (e.g.
CC   repeated transfusions). Data regarding the structure and function
CC   correlations of HFE polypeptides is useful in designing drugs that
CC   modulate the HFE gene and HFE activity. The polypeptides are also useful
CC   in protein replacement therapy for individuals possessing a defective
CC   HFE gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the
CC   polypeptides are also useful in treating primary and secondary iron
CC   overload diseases. The modulators of the transferrin receptor are useful
CC   in treating iron deficiency conditions such as anemia, and in modulating
CC   the amount of iron transported into a cell. The HFE polypeptides provide
CC   a molecular basis for the relationship between HFE and iron metabolism,
CC   which enables treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match             98.6%;  Score 1493;  DB 20;  Length 276;
  Best Local Similarity   98.9%;  Pred. No. 1e-130;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 10
ABG72686
ID   ABG72686 standard; protein; 276 AA.
XX
AC   ABG72686;
XX
DT   05-MAR-2003  (first entry)
XX
DE   Human haemochromatosis (HFE) mature protein, mutant H41D.
XX
KW   Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW   iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW   Beta2m; transferrin receptor; anaemia; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 41
FT                   /note= "Wild-type His substituted by Asp"
XX
PN   US6391852-B1.
XX
PD   21-MAY-2002.
XX
PF   12-JUN-1998;   98US-0094964.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (BIRA ) BIO-RAD LAB INC.
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI   Feder JN,  Bjorkman PJ,  Schatzman RC;
XX
DR   WPI; 2003-155377/15.
XX
PT   Method of treating an iron overload disease comprises administration of
PT   a soluble complex comprising a 276 amino acid HFE polypeptide and a
PT   full length, wild-type human beta2m -
XX
PS   Claim 2; Column 2; 17pp; English.
XX
CC   The invention relates to a method of treating an iron overload disease
CC   comprising administration of a soluble complex comprising a 276 amino
CC   acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC   (ABG72685-ABG72687) and a full length, wild-type human beta2m
CC   (beta2-microglobulin). In a HeLa cell based assay, binding and uptake of
CC   ^1^2^5I-transferrin in the presence of purified H63D-HFE/beta2m
CC   heterodimers was determined. At a concentration of 250 nM H63D-HFE/
CC   beta2m heterodimers, the transferrin receptor (TfR) displayed a KD for
CC   transferrin of 28 nM. At the same concentration of normal HFE/beta_2m
CC   heterodimers, TfR displayed a KD for transferrin of 40 nM. In the absence
CC   of any HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of
CC   7nM. It was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC   efficient in decreasing TfR affinity for transferrin compared to
CC   wild-type HFE. The method is useful for treating iron overload diseases
CC   and iron deficiency e.g. anaemia. The present sequence is the H63D
CC   (residue 63 of the full length protein, 41 of the mature form)
CC   mutant from of mature HFE used to investigate the role of the His
CC   residue in transferrin receptor binding to transferrin.
XX
SQ   Sequence   276 AA;

  Query Match             98.6%;  Score 1493;  DB 24;  Length 276;
  Best Local Similarity   98.9%;  Pred. No. 1e-130;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 11
AAB36871
ID   AAB36871 standard; Protein; 348 AA.
XX
AC   AAB36871;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
```

```
KW  T-cell differentiation factor; iron overload.
XX
OS  Homo sapiens.
XX
PN  US6140305-A.
XX
PD  31-OCT-2000.
XX
PF  04-APR-1997;  97US-0834497.
XX
PR  04-APR-1996;  96US-0630912.
PR  16-APR-1996;  96US-0632673.
PR  23-MAY-1996;  96US-0652265.
XX
PA  (BIRA ) BIO-RAD LAB INC.
XX
PI  Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI  Feder JN;
XX
DR  WPI; 2001-006341/01.
DR  N-PSDB; AAC68427.
XX
PT  New hereditary hemochromatosis gene products or polypeptides, useful
PT  for treating hereditary hemochromatosis in a patient, and as a metal
PT  chelation agent alleviating iron overload -
XX
PS  Claim 3; Fig 4; 108pp; English.
XX
CC  The present invention relates to hereditary hemochromatosis gene
CC  products. These proteins may be used to treat a patient diagnosed as
CC  having human hemochromatosis disease. It is also useful as a metal
CC  chelation agent or as a T-cell differentiation factor, and for
CC  alleviating iron overload. They may also be used in protein replacement
CC  therapy for individuals having a defective human hemochromatosis gene.
XX
SQ  Sequence  348 AA;
```

```
  Query Match            98.6%;  Score 1493;  DB 22;  Length 348;
  Best Local Similarity  98.9%;  Pred. No. 1.4e-130;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298
```

```
RESULT 12
AAB36870
ID  AAB36870 standard; Protein; 348 AA.
XX
AC  AAB36870;
XX
DT  21-FEB-2001  (first entry)
XX
DE  Human hereditary hemochromatosis 24d1 mutation protein.
XX
KW  HH; hereditary hemochromatosis; chelation agent;
```

```
KW  T-cell differentiation factor; iron overload.
XX
OS  Homo sapiens.
XX
PN  US6140305-A.
XX
PD  31-OCT-2000.
XX
PF  04-APR-1997;  97US-0834497.
XX
PR  04-APR-1996;  96US-0630912.
PR  16-APR-1996;  96US-0632673.
PR  23-MAY-1996;  96US-0652265.
XX
PA  (BIRA ) BIO-RAD LAB INC.
XX
PI  Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI  Feder JN;
XX
DR  WPI; 2001-006341/01.
DR  N-PSDB; AAC68426.
XX
PT  New hereditary hemochromatosis gene products or polypeptides, useful
PT  for treating hereditary hemochromatosis in a patient, and as a metal
PT  chelation agent alleviating iron overload -
XX
PS  Claim 2; Fig 3; 108pp; English.
XX
CC  The present invention relates to hereditary hemochromatosis gene
CC  products. These proteins may be used to treat a patient diagnosed as
CC  having human hemochromatosis disease. It is also useful as a metal
CC  chelation agent or as a T-cell differentiation factor, and for
CC  alleviating iron overload. They may also be used in protein replacement
CC  therapy for individuals having a defective human hemochromatosis gene.
XX
SQ  Sequence  348 AA;
```

```
  Query Match            98.5%;  Score 1491;  DB 22;  Length 348;
  Best Local Similarity  98.9%;  Pred. No. 2.1e-130;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

```
RESULT 13
AAB36872
ID  AAB36872 standard; Protein; 348 AA.
XX
AC  AAB36872;
XX
DT  21-FEB-2001  (first entry)
XX
DE  Human hereditary hemochromatosis 24d1/2 mutation protein.
XX
KW  HH; hereditary hemochromatosis; chelation agent;
```

```
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68428.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 4; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match             97.9%;  Score 1482;  DB 22;  Length 348;
  Best Local Similarity   98.6%;  Pred. No. 1.5e-129;
  Matches  272;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 14
AAB36873
ID   AAB36873 standard; Protein; 361 AA.
XX
AC   AAB36873;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Rabbit leukocyte antigen.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Oryctolagus cuniculus.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 7; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   361 AA;

  Query Match             34.1%;  Score 517;  DB 22;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 1.1e-39;
  Matches  111;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              | :    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   ||||  |    |||| || :||  |:: | :| :   :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 15
AAB36874
ID   AAB36874 standard; Protein; 365 AA.
XX
AC   AAB36874;
XX
DT   21-FEB-2001  (first entry)
XX
DE   MHC class I protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
```

XX
OS Unidentified.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 7; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 365 AA;

Query Match 33.6%; Score 508; DB 22; Length 365;
Best Local Similarity 39.7%; Pred. No. 7.5e-39;
Matches 110; Conservative 44; Mismatches 115; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
              :: :|      ||  |:   :| |:  |||||:: ||::  |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTLQMMFGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299
```

Search completed: August 5, 2003, 13:08:24
Job time : 39 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:10:04 ; Search time 33 Seconds
(without alignments)
993.264 Million cell updates/sec

Title: US-10-092-404-3
Perfect score: 1514
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1514	100.0	276	15	US-10-092-404-3	Sequence 3, Appli
2	1502	99.2	276	15	US-10-092-404-1	Sequence 1, Appli
3	1502	99.2	348	12	US-09-981-606-2	Sequence 2, Appli
4	1493	98.6	276	15	US-10-092-404-2	Sequence 2, Appli
5	504	33.3	92	14	US-10-016-634A-120	Sequence 120, App
6	500	33.0	280	15	US-10-073-300-6	Sequence 6, Appli
7	500	33.0	415	15	US-10-073-300-5	Sequence 5, Appli
8	486	32.1	298	15	US-10-205-823-40	Sequence 40, Appl
9	486	32.1	298	15	US-10-205-823-42	Sequence 42, Appl
10	486	32.1	298	15	US-10-177-293-23	Sequence 23, Appl
11	471	31.1	542	15	US-10-015-535-32	Sequence 32, Appl
12	471	31.1	542	15	US-10-015-535-34	Sequence 34, Appl
13	470	31.0	542	15	US-10-015-535-36	Sequence 36, Appl
14	468	30.9	540	15	US-10-015-535-22	Sequence 22, Appl
15	468	30.9	541	15	US-10-015-535-28	Sequence 28, Appl
16	468	30.9	542	15	US-10-015-535-24	Sequence 24, Appl
17	468	30.9	542	15	US-10-015-535-26	Sequence 26, Appl
18	446	29.5	332	9	US-09-870-521-3	Sequence 3, Appli
19	441	29.1	334	9	US-09-870-521-4	Sequence 4, Appli
20	439	29.0	540	15	US-10-015-535-30	Sequence 30, Appl
21	354.5	23.4	170	9	US-09-925-301-1307	Sequence 1307, Ap
22	330	21.8	271	9	US-09-925-301-1431	Sequence 1431, Ap
23	275	18.2	145	9	US-09-810-560-8	Sequence 8, Appli
24	273	18.0	181	11	US-09-013-077A-13	Sequence 13, Appl
25	237	15.7	184	10	US-09-858-580-21	Sequence 21, Appl
26	237	15.7	184	11	US-09-847-172-21	Sequence 21, Appl
27	226	14.9	91	9	US-09-864-761-38005	Sequence 38005, A
28	223	14.7	91	9	US-09-864-761-35461	Sequence 35461, A
29	217.5	14.4	171	15	US-10-144-929-116	Sequence 116, App
30	210.5	13.9	104	9	US-09-925-302-835	Sequence 835, App
31	207	13.7	117	9	US-09-810-560-9	Sequence 9, Appli
32	202.5	13.4	183	15	US-10-036-542-62	Sequence 62, Appl
33	196.5	13.0	93	9	US-09-864-761-39479	Sequence 39479, A
34	196.5	13.0	110	9	US-09-864-761-35339	Sequence 35339, A
35	196.5	13.0	114	9	US-09-864-761-37988	Sequence 37988, A
36	176.5	11.7	261	10	US-09-925-664-30	Sequence 30, Appl
37	173	11.4	110	10	US-09-796-692-799	Sequence 799, App
38	173	11.4	110	10	US-09-796-692-2139	Sequence 2139, Ap
39	173	11.4	110	15	US-10-040-862-799	Sequence 799, App
40	173	11.4	110	15	US-10-040-862-2139	Sequence 2139, Ap
41	170	11.2	411	14	US-10-015-536-17	Sequence 17, Appl
42	167.5	11.1	285	10	US-09-756-983-24	Sequence 24, Appl
43	167	11.0	246	9	US-09-989-722-225	Sequence 225, App
44	167	11.0	246	9	US-09-989-723-225	Sequence 225, App
45	167	11.0	246	9	US-09-989-279-225	Sequence 225, App

ALIGNMENTS

RESULT 1
US-10-092-404-3
; Sequence 3, Application US/10092404
; Publication No. US20030073627A1
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; Bjorkman, Pamela J.
; Schatzman, Randall C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
; AND IRON DEFICIENCY DISEASES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/092,404
; FILING DATE: 04-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,964
; FILING DATE: June 12, 1998
; APPLICATION NUMBER: 08/876,010
; FILING DATE: June 13, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0074-999
; TELECOMMUNICATION INFORMATION:

```
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-092-404-3

  Query Match            100.0%;  Score 1514;  DB 15;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-147;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-10-092-404-1
; Sequence 1, Application US/10092404
; Publication No. US20030073627A1
;   GENERAL INFORMATION:
;        APPLICANT: Feder, John N.
;                   Bjorkman, Pamela J.
;                   Schatzman, Randall C.
;        TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
;                            DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;                            AND IRON DEFICIENCY DISEASES
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Pennie & Edmonds, LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036-2811
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/092,404
;             FILING DATE: 04-Mar-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/094,964
;             FILING DATE: June 12, 1998
;             APPLICATION NUMBER: 08/876,010
;             FILING DATE: June 13, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Poissant, Brian M
;             REGISTRATION NUMBER: 28,462
;             REFERENCE/DOCKET NUMBER: 8907-0074-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-092-404-1

  Query Match            99.2%;  Score 1502;  DB 15;  Length 276;
  Best Local Similarity   99.3%;  Pred. No. 3.1e-146;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 3
US-09-981-606-2
; Sequence 2, Application US/09981606
; Publication No. US20030129595A1
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg et al.
;  TITLE OF INVENTION: Mutations associated with iron disorders
;  FILE REFERENCE: 24065-004CON
;  CURRENT APPLICATION NUMBER: US/09/981,606
;  CURRENT FILING DATE:  2002-10-16
;  PRIOR APPLICATION NUMBER: 09/277,457
;  PRIOR FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-981-606-2

  Query Match            99.2%;  Score 1502;  DB 12;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 4.3e-146;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142
```

```
Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
US-10-092-404-2
; Sequence 2, Application US/10092404
; Publication No. US20030073627A1
;   GENERAL INFORMATION:
;        APPLICANT: Feder, John N.
;                   Bjorkman, Pamela J.
;                   Schatzman, Randall C.
;        TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
;                            DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;                            AND IRON DEFICIENCY DISEASES
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Pennie & Edmonds, LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036-2811
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/092,404
;             FILING DATE: 04-Mar-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/094,964
;             FILING DATE: June 12, 1998
;             APPLICATION NUMBER: 08/876,010
;             FILING DATE: June 13, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Poissant, Brian M
;             REGISTRATION NUMBER: 28,462
;             REFERENCE/DOCKET NUMBER: 8907-0074-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-092-404-2

  Query Match             98.6%;  Score 1493;  DB 15;  Length 276;
  Best Local Similarity   98.9%;  Pred. No. 2.7e-145;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 5
US-10-016-634A-120
; Sequence 120, Application US/10016634A
; Publication No. US20020192666A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Ghosh, Malavika
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pr
;  FILE REFERENCE: DEX-0255
;  CURRENT APPLICATION NUMBER: US/10/016,634A
;  CURRENT FILING DATE:  2001-10-31
;  PRIOR APPLICATION NUMBER: US 60/244,258
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 176
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
;   LENGTH: 92
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-016-634A-120

  Query Match             33.3%;  Score 504;  DB 14;  Length 92;
  Best Local Similarity   98.9%;  Pred. No. 3.1e-44;
  Matches   91;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALEFCPDTLDWRAAEPRAWPTKLEWERHK 151
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 60

Qy        152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              ||||||||||||||||||||||||||||||||
Db         61 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 92


RESULT 6
US-10-073-300-6
; Sequence 6, Application US/10073300
; Publication No. US20030003535A1
; GENERAL INFORMATION:
;  APPLICANT: Reiter, Yoram
;  TITLE OF INVENTION: SINGLE CHAIN CLASS I MAJOR HISTO- COMPATIBILITY COMPLEXES
;  FILE REFERENCE: 02/23339
;  CURRENT APPLICATION NUMBER: US/10/073,300
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-073-300-6

  Query Match             33.0%;  Score 500;  DB 15;  Length 280;
  Best Local Similarity   39.4%;  Pred. No. 3.6e-43;
  Matches  109;  Conservative   44;  Mismatches  116;  Indels    8;  Gaps    7;
```

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db          2 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 60

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db         61 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        121 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 179

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        180 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 238

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        239 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 275


RESULT 7
US-10-073-300-5
; Sequence 5, Application US/10073300
; Publication No. US20030003535A1
; GENERAL INFORMATION:
;  APPLICANT: Reiter, Yoram
;  TITLE OF INVENTION: SINGLE CHAIN CLASS I MAJOR HISTO- COMPATIBILITY COMPLEXES
;  FILE REFERENCE: 02/23339
;  CURRENT APPLICATION NUMBER: US/10/073,300
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human beta2 microglobulin linked to MHC class I heavy chain
US-10-073-300-5

  Query Match             33.0%;  Score 500;  DB 15;  Length 415;
  Best Local Similarity   39.4%;  Pred. No. 6.2e-43;
  Matches  109;  Conservative   44;  Mismatches  116;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db        117 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 175

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNASKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
                :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db        176 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 235

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :| :    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        236 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 294

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        295 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 353

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||| |:| || :|| : ||
Db        354 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 390


RESULT 8
US-10-205-823-40
; Sequence 40, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Anderson, Dustin
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF PROSTATE CANCER
;  FILE REFERENCE: MRI-044
;  CURRENT APPLICATION NUMBER: US/10/205,823
;  CURRENT FILING DATE:  2002-07-25
;  PRIOR APPLICATION NUMBER: 60/307,982
;  PRIOR FILING DATE: 2001-07-25
;  PRIOR APPLICATION NUMBER: 60/314,356
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/325,020
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: 60/341,746
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-205-823-40

  Query Match             32.1%;  Score 486;  DB 15;  Length 298;
  Best Local Similarity   36.7%;  Pred. No. 1.1e-41;
  Matches  101;  Conservative   52;  Mismatches  112;  Indels   10;  Gaps    4;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQL 65
              :|| |:: | |:    :  |:||| ::|  |  |: : |: :|   |       : | |
Db         28 YSLTYIYTGLSKHVEDVPAFQALGSLNDLQFFRYNSKDRKSQPMGLWRQVE-GMEDWKQD 86

Qy         66 SQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALE 125
              ||  |  : :|       |:| :| |  || ||   |||:: : |:  :||| |||:| :|
Db         87 SQLQKAREDIFMETLKDIVEYYNDSNGSHVLQGRFGCEIENNRSSGAFWKYYYDGKDYIE 146

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              |  :   |   :| |  || :||   :  :: :|||| :||| |:: |:  : :||:| |
Db        147 FNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKYLKYSKNILDRQDP 206

Qy        186 PLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPK---DVLPNGDGT 241
              | | || |         |:| | ::||  | : | :  :        ||:   ||| ||:||
Db        207 PSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EPELRGDVLHNGNGT 261

Qy        242 YQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              || |: :|||| :   |:| |:|  | |||:| ||
Db        262 YQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 296


RESULT 9
US-10-205-823-42
; Sequence 42, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
```

```
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Anderson, Dustin
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF PROSTATE CANCER
;  FILE REFERENCE: MRI-044
;  CURRENT APPLICATION NUMBER: US/10/205,823
;  CURRENT FILING DATE:  2002-07-25
;  PRIOR APPLICATION NUMBER: 60/307,982
;  PRIOR FILING DATE: 2001-07-25
;  PRIOR APPLICATION NUMBER: 60/314,356
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/325,020
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: 60/341,746
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-205-823-42

  Query Match            32.1%;  Score 486;  DB 15;  Length 298;
  Best Local Similarity  36.7%;  Pred. No. 1.1e-41;
  Matches  101;  Conservative   52;  Mismatches  112;  Indels   10;  Gaps    4;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQL 65
              :|| |:: | |:    :  |:||| ::|  |  |: : |: :|   |      : | |
Db         28 YSLTYIYTGLSKHVEDVPAFQALGSLNDLQFFRYNSKDRKSQPMGLWRQVE-GMEDWKQD 86

Qy         66 SQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALE 125
              ||  |  : :|      |:| :| |  || ||   |||:: : |:  :||| |||:| :|
Db         87 SQLQKAREDIFMETLKDIVEYYNDSNGSHVLQGRFGCEIENNRSSGAFWKYYYDGKDYIE 146

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              |  :   |  :| | || :||   :  :: :|||| :||| |:: |:  : :||:| |
Db        147 FNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKYLKYSKNILDRQDP 206

Qy        186 PLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPK---DVLPNGDGT 241
              | | || |      |:| | ::||  | : | :  :       ||:   ||| ||:||
Db        207 PSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EPELRGDVLHNGNGT 261

Qy        242 YQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              || |: :|||| :   |:| |:|  | |||:| ||
Db        262 YQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 296


RESULT 10
US-10-177-293-23
; Sequence 23, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Gannavarpu, Manjula
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Mertens, Maureen
;  APPLICANT:  Myer, Vic
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Monahan, John
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  Bast Jr., Robert C.
;  APPLICANT:  Hortobagyi, Gabriel N.
;  APPLICANT:  Pusztai, Lajos
;  APPLICANT:  Meric, Funda
;  APPLICANT:  Sahin, Aysegul
;  APPLICANT:  Mills, Gordon B.
;  TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
;  TITLE OF INVENTION:  PREVENTION, AND THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-038
;  CURRENT APPLICATION NUMBER: US/10/177,293
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/299,887
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/301,572
;  PRIOR FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: US 60/306,501
;  PRIOR FILING DATE: 2001-07-18
;  PRIOR APPLICATION NUMBER: US 60/325,002
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US 60/362,585
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/xxx,xxx
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 506
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-177-293-23

  Query Match            32.1%;  Score 486;  DB 15;  Length 298;
  Best Local Similarity  36.7%;  Pred. No. 1.1e-41;
  Matches  101;  Conservative   52;  Mismatches  112;  Indels   10;  Gaps    4;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQL 65
              :|| |:: | |:    :  |:||| ::|  |  |: : |: :|   |      : | |
Db         28 YSLTYIYTGLSKHVEDVPAFQALGSLNDLQFFRYNSKDRKSQPMGLWRQVE-GMEDWKQD 86

Qy         66 SQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDALE 125
              ||  |  : :|      |:| :| |  || ||   |||:: : |:  :||| |||:| :|
Db         87 SQLQKAREDIFMETLKDIVEYYNDSNGSHVLQGRFGCEIENNRSSGAFWKYYYDGKDYIE 146

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
              |  :   |  :| | || :||   :  :: :|||| :||| |:: |:  : :||:| |
Db        147 FNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKYLKYSKNILDRQDP 206

Qy        186 PLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPK---DVLPNGDGT 241
              | | || |      |:| | ::||  | : | :  :       ||:   ||| ||:||
Db        207 PSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EPELRGDVLHNGNGT 261

Qy        242 YQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              || |: :|||| :   |:| |:|  | |||:| ||
Db        262 YQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 296


RESULT 11
US-10-015-535-32
; Sequence 32, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
```

```
;   OTHER INFORMATION: peptide
US-10-015-535-32

  Query Match             31.1%;  Score 471;  DB 15;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 8.5e-40;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              | :    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 12
US-10-015-535-34
; Sequence 34, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-34

  Query Match             31.1%;  Score 471;  DB 15;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 8.5e-40;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              | :    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 13
US-10-015-535-36
; Sequence 36, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-36

  Query Match             31.0%;  Score 470;  DB 15;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 1.1e-39;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFY--DHESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :  | |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFNSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              | :    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 14
US-10-015-535-22
; Sequence 22, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 540
```

```
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-22

  Query Match             30.9%;  Score 468;  DB 15;  Length 540;
  Best Local Similarity   39.5%;  Pred. No. 1.7e-39;
  Matches  109;  Conservative   39;  Mismatches  120;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db        144 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 202

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        203 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 262

Qy        122 DALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              | :    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db        263 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 321

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        322 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 380

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        381 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 416


RESULT 15
US-10-015-535-28
; Sequence 28, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 541
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-28

  Query Match             30.9%;  Score 468;  DB 15;  Length 541;
  Best Local Similarity   39.5%;  Pred. No. 1.7e-39;
  Matches  109;  Conservative   39;  Mismatches  120;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
              ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db         56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNASK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
              : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db        115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy        122 DALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              | :    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db        175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 233

Qy        182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db        234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| | :: || |:|| ||| | | || :|| : ||
Db        293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


Search completed: August  5, 2003, 13:21:56
Job time : 34 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:07:04 ; Search time 14.5 Seconds
(without alignments)
805.365 Million cell updates/sec

Title: US-10-092-404-3
Perfect score: 1514
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1514	100.0	276	4	US-09-094-964-3	Sequence 3, Appli
2	1502	99.2	276	4	US-09-094-964-1	Sequence 1, Appli
3	1502	99.2	348	3	US-08-652-265-2	Sequence 2, Appli
4	1502	99.2	348	3	US-08-834-497A-2	Sequence 2, Appli
5	1502	99.2	348	3	US-09-503-444A-2	Sequence 2, Appli
6	1502	99.2	348	4	US-09-277-457-2	Sequence 2, Appli
7	1502	99.2	348	4	US-09-679-729-2	Sequence 2, Appli
8	1493	98.6	276	4	US-09-094-964-2	Sequence 2, Appli
9	1493	98.6	348	3	US-08-652-265-6	Sequence 6, Appli
10	1493	98.6	348	3	US-08-834-497A-6	Sequence 6, Appli
11	1493	98.6	348	3	US-09-503-444A-6	Sequence 6, Appli
12	1491	98.5	348	3	US-08-652-265-4	Sequence 4, Appli
13	1491	98.5	348	3	US-08-834-497A-4	Sequence 4, Appli
14	1491	98.5	348	3	US-09-503-444A-4	Sequence 4, Appli
15	1482	97.9	348	3	US-08-652-265-8	Sequence 8, Appli
16	1482	97.9	348	3	US-08-834-497A-8	Sequence 8, Appli
17	1482	97.9	348	3	US-09-503-444A-8	Sequence 8, Appli
18	517	34.1	361	3	US-08-652-265-22	Sequence 22, Appl
19	517	34.1	361	3	US-08-834-497A-22	Sequence 22, Appl
20	517	34.1	361	3	US-09-503-444A-22	Sequence 22, Appl
21	511	33.8	364	4	US-08-914-372C-11	Sequence 11, Appl
22	508	33.6	365	3	US-08-652-265-23	Sequence 23, Appl
23	508	33.6	365	3	US-08-834-497A-23	Sequence 23, Appl
24	508	33.6	365	3	US-09-503-444A-23	Sequence 23, Appl
25	500	33.0	274	2	US-08-484-905-107	Sequence 107, App
26	500	33.0	274	3	US-08-481-985B-107	Sequence 107, App
27	500	33.0	274	3	US-08-370-476-107	Sequence 107, App
28	500	33.0	341	3	US-08-890-719-38	Sequence 38, Appl
29	499	33.0	365	2	US-08-484-905-97	Sequence 97, Appl
30	499	33.0	365	3	US-08-481-985B-97	Sequence 97, Appl
31	499	33.0	365	3	US-08-370-476-97	Sequence 97, Appl
32	498	32.9	274	2	US-08-484-905-108	Sequence 108, App
33	498	32.9	274	3	US-08-481-985B-108	Sequence 108, App
34	498	32.9	274	3	US-08-370-476-108	Sequence 108, App
35	498	32.9	365	2	US-08-484-905-100	Sequence 100, App
36	498	32.9	365	3	US-08-481-985B-100	Sequence 100, App
37	498	32.9	365	3	US-08-370-476-100	Sequence 100, App
38	497	32.8	274	1	US-08-222-851-1	Sequence 1, Appli
39	497	32.8	363	4	US-08-914-372C-37	Sequence 37, Appl
40	497	32.8	365	2	US-08-484-905-99	Sequence 99, Appl
41	497	32.8	365	3	US-08-481-985B-99	Sequence 99, Appl
42	497	32.8	365	3	US-08-370-476-99	Sequence 99, Appl
43	496	32.8	274	2	US-08-484-905-106	Sequence 106, App
44	496	32.8	274	3	US-08-481-985B-106	Sequence 106, App
45	496	32.8	274	3	US-08-370-476-106	Sequence 106, App

ALIGNMENTS

```
RESULT 1
US-09-094-964-3
; Sequence 3, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-3

  Query Match           100.0%;  Score 1514;  DB 4;  Length 276;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-144;
```

```
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-09-094-964-1
; Sequence 1, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-1

  Query Match            99.2%;  Score 1502;  DB 4;  Length 276;
  Best Local Similarity  99.3%;  Pred. No. 9.9e-143;

  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 3
US-08-652-265-2
; Sequence 2, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-2

  Query Match            99.2%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity  99.3%;  Pred. No. 1.4e-142;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
                |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
                || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                ||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
US-08-834-497A-2
; Sequence 2, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-2

  Query Match             99.2%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-142;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
                |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
                || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                ||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 5
US-09-503-444A-2
; Sequence 2, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-2

  Query Match             99.2%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-142;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 6
US-09-277-457-2
; Sequence 2, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-277-457-2

  Query Match             99.2%;  Score 1502;  DB 4;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-142;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 7
US-09-679-729-2
; Sequence 2, Application US/09679729
; Patent No. 6509442
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 24065-004 DIV
;  CURRENT APPLICATION NUMBER: US/09/679,729
;  CURRENT FILING DATE:  2000-10-04
;  PRIOR APPLICATION NUMBER: 09/277,457
;  PRIOR FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-679-729-2

  Query Match             99.2%;  Score 1502;  DB 4;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-142;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMPTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 8
US-09-094-964-2
; Sequence 2, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
```

```
;       COUNTRY:  USA
;       ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows
;       SOFTWARE:  FastSEQ for Windows Version 2.0b
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/094,964
;       FILING DATE:  June 12, 1998
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/876,010
;       FILING DATE:  June 13, 1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Poissant, Brian M
;       REGISTRATION NUMBER:  28,462
;       REFERENCE/DOCKET NUMBER:  8907-0074-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  276 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-094-964-2

  Query Match             98.6%;  Score 1493;  DB 4;  Length 276;
  Best Local Similarity   98.9%;  Pred. No. 7.9e-142;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 9
US-08-652-265-6
; Sequence 6, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-6

  Query Match             98.6%;  Score 1493;  DB 3;  Length 348;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-141;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 10
US-08-834-497A-6
; Sequence 6, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
```

```
;       ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  FastSEQ for Windows Version 2.0b
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/834,497A
;       FILING DATE:  04-APR-1997
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/652,265
;       FILING DATE:  23-MAY-1996
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/632,673
;       FILING DATE:  16-APR-1996
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/630,912
;       FILING DATE:  04-APR-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Poissant, Brian M.
;       REGISTRATION NUMBER:  28,462
;       REFERENCE/DOCKET NUMBER:  8907-0056-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  348 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-834-497A-6

  Query Match             98.6%;  Score 1493;  DB 3;  Length 348;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-141;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 11
US-09-503-444A-6
; Sequence 6, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;     APPLICANT:  Thomas, Winston J.
;     APPLICANT:  Drayna, Dennis T.
;     APPLICANT:  Feder, John N.
;     APPLICANT:  Gnirke, Andreas
;     APPLICANT:  Ruddy, David
;     APPLICANT:  Tsuchihashi, Zenta
;     APPLICANT:  Wolff, Roger K.
;     TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;     NUMBER OF SEQUENCES:  44
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds LLP
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect Version 8
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/503,444A
;       FILING DATE:  14-Feb-2000
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/652,265
;       FILING DATE:  23-May-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/632,673
;       FILING DATE:  16-Apr-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/630,912
;       FILING DATE:  04-Apr-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Poissant, Brian M.
;       REGISTRATION NUMBER:  28,462
;       REFERENCE/DOCKET NUMBER:  8907-0088-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-790-9090
;       TELEFAX:  212-869-9741
;       TELEX:  66141
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  348 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-503-444A-6

  Query Match             98.6%;  Score 1493;  DB 3;  Length 348;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-141;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 12
US-08-652-265-4
; Sequence 4, Application US/08652265
```

```
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-4

  Query Match             98.5%;  Score 1491;  DB 3;  Length 348;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-141;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 13
US-08-834-497A-4
; Sequence 4, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-4

  Query Match             98.5%;  Score 1491;  DB 3;  Length 348;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-141;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

```
                  |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 14
US-09-503-444A-4
; Sequence 4, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-4

  Query Match             98.5%;  Score 1491;  DB 3;  Length 348;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-141;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 15
US-08-652-265-8
; Sequence 8, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-8

  Query Match             97.9%;  Score 1482;  DB 3;  Length 348;
  Best Local Similarity   98.6%;  Pred. No. 1.4e-140;
  Matches  272;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                ||||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

Search completed: August 5, 2003, 13:11:12
Job time : 15.5 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:05:29 ; Search time 38 Seconds
(without alignments)
1152.856 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1520	100.0	276	20	AAW94296	HFE mutant (H63D-H
2	1520	100.0	276	24	ABG72686	Human haemochromat
3	1520	100.0	348	22	AAB36871	Human hereditary h
4	1513	99.5	276	20	AAW94295	Wild-type HFE poly
5	1513	99.5	276	24	ABG72685	Human haemochromat
6	1513	99.5	348	18	AAW36499	Hereditary haemoch
7	1513	99.5	348	21	AAB19149	A human histocompa
8	1513	99.5	348	22	AAB36869	Human hereditary h
9	1509	99.3	348	22	AAB36872	Human hereditary h
10	1508	99.2	438	23	AAU80035	Beta 2 microglobul
11	1502	98.8	348	22	AAB36870	Human hereditary h
12	1493	98.2	276	20	AAW94297	HFE mutant (H111A/
13	1493	98.2	276	24	ABG72687	Human haemochromat
14	523	34.4	361	22	AAB36873	Rabbit leukocyte a
15	514	33.8	92	24	ABP68379	Human colon specif
16	514	33.8	365	22	AAB36874	MHC class I protei
17	506	33.3	274	21	AAY68275	Human leukocyte an
18	506	33.3	274	21	AAY52929	HLA-A2/A28 family
19	506	33.3	274	22	AAB58690	HLA-A2/A28 protein
20	506	33.3	280	22	AAU10225	Human leukocyte an
21	506	33.3	280	24	ABU08672	Human histocompati
22	506	33.3	415	22	AAU10224	Human partial beta
23	506	33.3	415	24	ABU08671	Human single chain
24	505	33.2	365	21	AAY68265	Human leukocyte an
25	505	33.2	365	21	AAY52919	HLA-A2/A28 family
26	505	33.2	365	22	AAB58680	HLA-A2/A28 protein
27	505	33.2	368	22	AAM24017	Human EST encoded
28	504	33.2	274	21	AAY68276	Human leukocyte an
29	504	33.2	274	21	AAY52930	HLA-A2/A28 family
30	504	33.2	274	22	AAB58691	HLA-A2/A28 protein
31	504	33.2	365	21	AAY68268	Human leukocyte an
32	504	33.2	365	21	AAY52922	HLA-A2/A28 family
33	504	33.2	365	22	AAB58683	HLA-A2/A28 protein
34	503	33.1	274	9	AAP80911	Consensus sequence
35	503	33.1	365	21	AAY68267	Human leukocyte an
36	503	33.1	365	21	AAY52921	HLA-A2/A28 family
37	503	33.1	365	22	AAB58682	HLA-A2/A28 protein
38	502	33.0	274	21	AAY68274	Human leukocyte an
39	502	33.0	274	21	AAY52928	HLA-A2/A28 family
40	502	33.0	274	22	AAB58689	HLA-A2/A28 protein
41	502	33.0	365	21	AAY68266	Human leukocyte an
42	502	33.0	365	21	AAY52920	HLA-A2/A28 family
43	502	33.0	365	22	AAB58681	HLA-A2/A28 protein
44	501	33.0	412	19	AAW68385	Chimeric HLA-A2.1/
45	500	32.9	274	21	AAY68273	Human leukocyte an

ALIGNMENTS

RESULT 1
AAW94296
ID AAW94296 standard; peptide; 276 AA.
XX
AC AAW94296;
XX
DT 27-APR-1999 (first entry)
XX
DE HFE mutant (H63D-HFE) polypeptide sequence.
XX
KW HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW transfusion; protein replacement therapy; hereditary hemochromatosis;
KW transferrin receptor; iron deficiency; anemia; mutant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "indicated in the sequence listing as Arg"
FT Misc-difference 41
FT /label= H63D
FT /note= "wild type His (of the mature protein sequence)
FT is replaced by Asp"
XX
PN WO9856814-A1.
XX
PD 17-DEC-1998.
XX
PF 12-JUN-1998; 98WO-US12436.
XX
PR 13-JUN-1997; 97US-0876010.
XX

PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (PROG-) PROGENITOR INC.
XX
PI Bjorkman PJ, Feder JN, Schatzman RC;
XX
DR WPI; 1999-080886/07.
XX
PT New treatment of an iron overload disease - comprises use of HFE
PT polypeptides provided in a complex with full length, wild type human
PT (2m), useful in protein replacement therapy
XX
PS Claim 3; Page 14; 36pp; English.
XX
CC The present sequence represents a H63D-HFE mutant polypeptide. The HFE
CC polypeptides (AAW94295-297) provided in a complex with full length,
CC wild type human beta-2-microglobulin (beta2m) form compositions in the
CC treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC other iron overload conditions resulting from secondary causes (e.g.
CC repeated transfusions). Data regarding the structure and function
CC correlations of HFE polypeptides is useful in designing drugs that
CC modulate the HFE gene and HFE activity. The polypeptides are also useful
CC in protein replacement therapy for individuals possessing a defective
CC HFE gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the
CC polypeptides are also useful in treating primary and secondary iron
CC overload diseases. The modulators of the transferrin receptor are useful
CC in treating iron deficiency conditions such as anemia, and in modulating
CC the amount of iron transported into a cell. The HFE polypeptides provide
CC a molecular basis for the relationship between HFE and iron metabolism,
CC which enables treatment of iron overload and deficiency diseases.
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1520; DB 20; Length 276;
Best Local Similarity 100.0%; Pred. No. 7.3e-135;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
Db 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
Db 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy 121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
Db 121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
Db 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy 241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db 241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276

RESULT 2
ABG72686
ID ABG72686 standard; protein; 276 AA.
XX
AC ABG72686;
XX
DT 05-MAR-2003 (first entry)
XX
DE Human haemochromatosis (HFE) mature protein, mutant H41D.
XX
KW Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW Beta2m; transferrin receptor; anaemia; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 41
FT /note= "Wild-type His substituted by Asp"
XX
PN US6391852-B1.
XX
PD 21-MAY-2002.
XX
PF 12-JUN-1998; 98US-0094964.
XX
PR 13-JUN-1997; 97US-0876010.
XX
PA (BIRA) BIO-RAD LAB INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 2003-155377/15.
XX
PT Method of treating an iron overload disease comprises administration of
PT a soluble complex comprising a 276 amino acid HFE polypeptide and a
PT full length, wild-type human beta2m -
XX
PS Claim 2; Column 2; 17pp; English.
XX
CC The invention relates to a method of treating an iron overload disease
CC comprising administration of a soluble complex comprising a 276 amino
CC acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC (ABG72685-ABG72687) and a full length, wild-type human beta2m
CC (beta2-microglobulin). In a HeLa cell based assay, binding and uptake of
CC ^1^2^5I-transferrin in the presence of purified H63D-HFE/beta2m
CC heterodimers was determined. At a concentration of 250 nM H63D-HFE/
CC beta2m heterodimers, the transferrin receptor (TfR) displayed a KD for
CC transferrin of 28 nM. At the same concentration of normal HFE/beta_2m
CC heterodimers, TfR displayed a KD for transferrin of 40 nM. In the absence
CC of any HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of
CC 7nM. It was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC efficient in decreasing TfR affinity for transferrin compared to
CC wild-type HFE. The method is useful for treating iron overload diseases
CC and iron deficiency e.g. anaemia. The present sequence is the H63D
CC (residue 63 of the full length protein, 41 of the mature form)
CC mutant from of mature HFE used to investigate the role of the His
CC residue in transferrin receptor binding to transferrin.
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1520; DB 24; Length 276;
Best Local Similarity 100.0%; Pred. No. 7.3e-135;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
Db 1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
Db 61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy 121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
Db 121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
Db 181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy 241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db 241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276

RESULT 3

```
AAB36871
ID   AAB36871 standard; Protein; 348 AA.
XX
AC   AAB36871;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68427.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 3; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match           100.0%;  Score 1520;  DB 22;  Length 348;
  Best Local Similarity  100.0%;  Pred. No. 9.9e-135;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
AAW94295
ID   AAW94295 standard; peptide; 276 AA.
XX
AC   AAW94295;
XX
DT   27-APR-1999  (first entry)
XX
DE   Wild-type HFE polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US12436.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (PROG-) PROGENITOR INC.
XX
PI   Bjorkman PJ,  Feder JN,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
XX
PS   Claim 1; Page 13; 36pp; English.
XX
CC   The present sequence represents a wild-type HFE polypeptide. The HFE
CC   polypeptides (AAW94295-297)  provided in a complex with full length,
CC   wild type human beta-2-microglobulin (beta2m) form compositions in the
CC   treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC   other iron overload conditions resulting from secondary causes (e.g.
CC   repeated transfusions). Data regarding the structure and function
CC   correlations of HFE polypeptides is useful in designing drugs that
CC   modulate the HFE gene and HFE activity. The polypeptides are also useful
CC   in protein replacement therapy for individuals possessing a defective
CC   HFE gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the
CC   polypeptides are also useful in treating primary and secondary iron
CC   overload diseases. The modulators of the transferrin receptor are useful
CC   in treating iron deficiency conditions such as anemia, and in modulating
CC   the amount of iron transported into a cell. The HFE polypeptides provide
CC   a molecular basis for the relationship between HFE and iron metabolism,
CC   which enables treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match           99.5%;  Score 1513;  DB 20;  Length 276;
  Best Local Similarity  99.6%;  Pred. No. 3.3e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 5
ABG72685
ID   ABG72685 standard; protein; 276 AA.
XX
AC   ABG72685;
XX
DT   05-MAR-2003  (first entry)
XX
DE   Human haemochromatosis (HFE) mature protein.
XX
KW   Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW   iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW   Beta2m; transferrin receptor; anaemia.
XX
OS   Homo sapiens.
XX
PN   US6391852-B1.
XX
PD   21-MAY-2002.
XX
PF   12-JUN-1998;   98US-0094964.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (BIRA ) BIO-RAD LAB INC.
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI   Feder JN,  Bjorkman PJ,  Schatzman RC;
XX
DR   WPI; 2003-155377/15.
XX
PT   Method of treating an iron overload disease comprises administration of
PT   a soluble complex comprising a 276 amino acid HFE polypeptide and a
PT   full length, wild-type human beta2m -
XX
PS   Claim 1; Column 1; 17pp; English.
XX
CC   The invention relates to a method of treating an iron overload disease
CC   comprising administration of a soluble complex comprising a 276 amino
CC   acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC   (ABG72685-ABG72687) and a full length, wild-type human beta2m
CC   (beta2-microglobulin). In a HeLa cell based assay, binding and uptake of
CC   ^1^2^5I-transferrin in the presence of purified H63D-HFE/beta2m
CC   heterodimers was determined. At a concentration of 250 nM H63D-HFE/
CC   beta2m heterodimers, the transferrin receptor (TfR) displayed a KD for
CC   transferrin of 28 nM. At the same concentration of normal HFE/beta_2m
CC   heterodimers, TfR displayed a KD for transferrin of 40 nM. In the absence
CC   of any HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of
CC   7nM. It was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC   efficient in decreasing TfR affinity for transferrin compared to
CC   wild-type HFE. The method is useful for treating iron overload diseases
CC   and iron deficiency e.g. anaemia. The present sequence is wild-type
CC   mature HFE.
XX
SQ   Sequence   276 AA;

  Query Match             99.5%;  Score 1513;  DB 24;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||

Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 6
AAW36499
ID   AAW36499 standard; Protein; 348 AA.
XX
AC   AAW36499;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene product.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 63
FT                   /note= "substituted by Asp in 24s2 mutant"
FT   Misc-difference 65
FT                   /note= "substituted by Cys in 24d7 variant"
FT   Misc-difference 282
FT                   /note= "substituted by Tyr in 24d1 mutant"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   N-PSDB; AAT96690, AAT96691.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 4; 115pp; English.
XX
CC   This polypeptide is the expression product of a novel human gene
CC   (see AAT96690) whose mutated form is associated with hereditary
CC   haemochromatosis (HH).  A single mutation (24d1) in the HH gene
CC   appears responsible for the majority of HH disease.  This comprises
CC   a G to A substitution that is present in 86% of affected
CC   chromosomes and in 4% of unaffected chromosomes.  It results in a
CC   Cys to Tyr substitution in the encoded protein at a critical
CC   disulphide bridge important for secondary structure.  The following
```

```
CC   are claimed: the 10825 bp genomic DNA sequence (I), a 1437 bp cDNA
CC   sequence (Ia) (see AAT96691) and their 24d1, 24d2 and 24d7 variants;
CC   a cloning or expression vector; host cells; a peptide product
CC   chosen from the HH gene product, its variants (24d1, 24d2 and
CC   24d7), or a peptide of at least 56 amino acid residues of these; an
CC   antibody produced using the peptide as an immunogen; a method to
CC   determine the presence or absence of the common HH gene mutation;
CC   an animal model for the HH disease; metal chelation agents, T-cell
CC   differentiation factors and therapeutic agents for the mitigation
CC   of injury due to oxidative process in vivo or mitigation of iron
CC   overload; a method for screening potential therapeutic agents for
CC   activity in connection with HH disease; an antisense oligonucleotide
CC   directed against a transcriptional product of a nucleic acid
CC   sequence as above; and oligonucleotides or pairs of oligonucleotides
CC   covering a range of nucleotides from (I), (Ia) or their variants,
CC   useful for detecting a polymorphism in the HH gene.  The invention
CC   also relates to methods for screening for HH homozygotes, to HH
CC   diagnosis, prenatal screening and diagnosis, and therapies of HH
CC   disease, including gene therapy, protein- and antibody-based
CC   therapeutics, and small molecule therapeutics.
XX
SQ   Sequence   348 AA;

  Query Match             99.5%;  Score 1513;  DB 18;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 7
AAB19149
ID   AAB19149 standard; Protein; 348 AA.
XX
AC   AAB19149;
XX
DT   19-FEB-2001  (first entry)
XX
DE   A human histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /note= "signal peptide"
FT   Misc-difference 63
FT                   /note= "when nucleotide 187 is mutated to G, then
FT                         this residue is Asp"
FT   Misc-difference 65
FT                   /note= "when nucleotide 193 is mutated to T, then
FT                         this residue is Cys"
FT   Domain          80..108
FT                   /note= "alpha1 domain"
FT   Misc-difference 93
FT                   /note= "when nucleotide 277 is mutated to C, then
FT                         this residue is Arg"
FT   Misc-difference 105
FT                   /note= "when nucleotide 314 is mutated to C, then
FT                         this residue is Thr"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   N-PSDB; AAA96769.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Disclosure; Page 3; 55pp; English.
XX
CC   The present sequence represents a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence   348 AA;

  Query Match             99.5%;  Score 1513;  DB 21;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 8
AAB36869
```

```
ID   AAB36869 standard; Protein; 348 AA.
XX
AC   AAB36869;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68425.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 1; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;
```

```
  Query Match            99.5%;  Score 1513;  DB 22;  Length 348;
  Best Local Similarity  99.6%;  Pred. No. 4.5e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298
```

RESULT 9
AAB36872

```
ID   AAB36872 standard; Protein; 348 AA.
XX
AC   AAB36872;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1/2 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68428.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 4; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;
```

```
  Query Match            99.3%;  Score 1509;  DB 22;  Length 348;
  Best Local Similarity  99.6%;  Pred. No. 1.1e-133;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

RESULT 10
AAU80035

ID AAU80035 standard; Protein; 438 AA.
XX
AC AAU80035;
XX
DT 15-JUL-2002 (first entry)
XX
DE Beta 2 microglobulin (beta2M)/HFE monochain.
XX
KW Human; beta 2 microglobulin; beta2M/HFE monochain; HFE; ischaemia;
KW iron absorption regulator; intracellular iron absorption; lung injury;
KW haemochromatosis; transfusion; thalassaemia; haemolytic anaemia;
KW chronic infection; transferrin receptor; TfR; brain tumour; cancer;
KW oxidative stress disorder; tissue damage; vascular disease;
KW inflammation; atherosclerosis; autoimmune disease;
KW inflammatory condition.
XX
OS Homo sapiens.
XX
PN WO200224929-A2.
XX
PD 28-MAR-2002.
XX
PF 24-SEP-2001; 2001WO-US29873.
XX
PR 22-SEP-2000; 2000US-234843P.
XX
PA (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
PA (MCIN/) MCINNIS P.
XX
PI Ehrlich R, Rotem-Yehudar R, Laham N;
XX
DR WPI; 2002-383192/41.
DR N-PSDB; ABK49917.
XX
PT Soluble beta 2 microglobulin/HFE monochain useful for treating
PT iron-overload conditions e.g. thalassemia and chronic infections,
PT comprises human beta 2 microglobulin linked to alpha domains of HFE by
PT a linker peptide -
XX
PS Example 2; Fig 2; 77pp; English.
XX
CC The invention relates to a soluble polypeptide (I) of beta 2
CC microglobulin (beta2m)/HFE monochain comprising human beta2m (or its
CC analogue or active fragment), linked to alpha1-alpha3 domains of human
CC HFE (a central regulator of iron absorption; undefined), or its analogue
CC or active fragment, by a flexible linker peptide, or a functional
CC derivative or salt of (I). (I) is useful for reducing intracellular iron
CC absorption in patients having hereditary haemochromatosis, transfusions,
CC thalassaemias, haemolytic anaemia or chronic infections, and for
CC delivering a therapeutic to cells that over-express transferrin receptor
CC (TfR) which are preferably lymphocytes or leukocytes, across the blood-
CC brain barrier. (I) is further useful for treating brain tumour. (I)
CC is also useful for treating oxidative stress disorders resulting in
CC tissue damage e.g. vascular diseases, inflammation, atherosclerosis,
CC lung injury, ischaemia, etc. A DNA molecule (II) encoding (I) is useful
CC as a platform for drug delivery of therapeutic use for cancer,
CC autoimmune diseases and inflammatory conditions. The monochain manifests
CC specific characteristics advantageous for drug delivery systems. It is a
CC soluble, stable and fully conformed protein. It binds specifically to
CC transferrin receptor (TfR) and therefore targets cells that over-express
CC this receptor. It is continuously internalised by the target cells, thus
CC enabling efficient drug delivery. It dissociates from the receptor in the
CC cells, minimising side effects. It negatively regulates iron absorption,
CC reducing growth of undesired cells and preventing lymphocyte activation.
CC It is not diluted in the blood as is transferrin. It should not induce an
CC immune response since it is a self non-polymeric protein and delivery of
CC drugs via monochain is expected to overcome drug-resistance since it is a
CC natural TfR-binding protein. The present sequence represents the amino
CC acid sequence of beta2m/HFE monochain.
XX
SQ Sequence 438 AA;

Query Match 99.2%; Score 1508; DB 23; Length 438;
Best Local Similarity 99.6%; Pred. No. 1.8e-133;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        135 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 194

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 254

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 314

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 374

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
              ||||||||||||||||||||||||||||||||||
Db        375 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 409
```

RESULT 11
AAB36870
ID AAB36870 standard; Protein; 348 AA.
XX
AC AAB36870;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation protein.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR N-PSDB; AAC68426.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Claim 2; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 348 AA;

Query Match 98.8%; Score 1502; DB 22; Length 348;

```
  Best Local Similarity   99.3%;  Pred. No. 4.9e-133;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 12
AAW94297
ID   AAW94297 standard; peptide; 276 AA.
XX
AC   AAW94297;
XX
DT   27-APR-1999  (first entry)
XX
DE   HFE mutant (H111A/H145A-HFE) polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia; mutant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
FT   Misc-difference 89
FT                   /label= H111A
FT                   /note= "wild type His (of the mature protein sequence)
FT                          is replaced by Ala"
FT   Misc-difference 123
FT                   /label= H145A
FT                   /note= "wild type His (of the mature protein sequence)
FT                          is replaced by Ala"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US12436.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (PROG-) PROGENITOR INC.
XX
PI   Bjorkman PJ,  Feder JN,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
XX
PS   Claim 5; Page 15; 36pp; English.
XX
CC   The present sequence represents a H111A/H145A-HFE mutant polypeptide.
CC   The HFE polypeptides (AAW94295-297) provided in a complex with full
CC   length, wild type human beta-2-microglobulin (beta2m) form compositions
CC   in the treatment of primary iron overload diseases (e.g.
CC   hemochromatosis), or other iron overload conditions resulting from
CC   secondary causes (e.g. repeated transfusions). Data regarding the
CC   structure and function correlations of HFE polypeptides is useful in
CC   designing drugs that modulate the HFE gene and HFE activity. The
CC   polypeptides are also useful in protein replacement therapy for
CC   individuals possessing a defective HFE gene (e.g. Hereditary
CC   hemochromatosis). (Ant)agonists of the polypeptides are also useful in
CC   treating primary and secondary iron overload diseases. The modulators of
CC   the transferrin receptor are useful in treating iron deficiency
CC   conditions such as anemia, and in modulating the amount of iron
CC   transported into a cell. The HFE polypeptides provide a molecular basis
CC   for the relationship between HFE and iron metabolism, which enables
CC   treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match             98.2%;  Score 1493;  DB 20;  Length 276;
  Best Local Similarity   98.9%;  Pred. No. 2.5e-132;
  Matches  273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 13
ABG72687
ID   ABG72687 standard; protein; 276 AA.
XX
AC   ABG72687;
XX
DT   05-MAR-2003  (first entry)
XX
DE   Human haemochromatosis (HFE) mature protein, mutant H89A/H123A.
XX
KW   Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW   iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW   Beta2m; transferrin receptor; anaemia; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 89
FT                   /note= "Wild-type His substituted by Ala"
FT   Misc-difference 123
FT                   /note= "Wild-type His substituted by Ala"
XX
PN   US6391852-B1.
XX
PD   21-MAY-2002.
XX
PF   12-JUN-1998;   98US-0094964.
```

XX
PR 13-JUN-1997; 97US-0876010.
XX
PA (BIRA) BIO-RAD LAB INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 2003-155377/15.
XX
PT Method of treating an iron overload disease comprises administration of
PT a soluble complex comprising a 276 amino acid HFE polypeptide and a
PT full length, wild-type human beta2m -
XX
PS Claim 3; Column 2; 17pp; English.
XX
CC The invention relates to a method of treating an iron overload disease
CC comprising administration of a soluble complex comprising a 276 amino
CC acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC (ABG72685-ABG72687) and a full length, wild-type human beta2m
CC (beta2-microglobulin). In a HeLa cell based assay, binding and uptake of
CC ^1^2^5I-transferrin in the presence of purified H63D-HFE/beta2m
CC heterodimers was determined. At a concentration of 250 nM H63D-HFE/
CC beta2m heterodimers, the transferrin receptor (TfR) displayed a KD for
CC transferrin of 28 nM. At the same concentration of normal HFE/beta_2m
CC heterodimers, TfR displayed a KD for transferrin of 40 nM. In the absence
CC of any HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of
CC 7nM. It was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC efficient in decreasing TfR affinity for transferrin compared to
CC wild-type HFE. The method is useful for treating iron overload diseases
CC and iron deficiency e.g. anaemia. The present sequence is the H111A/H145A
CC (residues 111 and 145 of the full length protein, 89/123 of the mature
CC form) mutant from of mature HFE used to investigate the role of the His
CC residues in transferrin receptor binding to transferrin.
XX
SQ Sequence 276 AA;

Query Match 98.2%; Score 1493; DB 24; Length 276;
Best Local Similarity 98.9%; Pred. No. 2.5e-132;
Matches 273; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

RESULT 14
AAB36873
ID AAB36873 standard; Protein; 361 AA.
XX
AC AAB36873;
XX
DT 21-FEB-2001 (first entry)
XX
DE Rabbit leukocyte antigen.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload.
XX
OS Oryctolagus cuniculus.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 7; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 361 AA;

Query Match 34.4%; Score 523; DB 22; Length 361;
Best Local Similarity 40.1%; Pred. No. 9.1e-41;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   | :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |    | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::    |   | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |    |||| || :||  |:: | :| :   :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 15
ABP68379
ID ABP68379 standard; Protein; 92 AA.
XX
AC ABP68379;
XX
DT 08-JAN-2003 (first entry)
XX
DE Human colon specific protein, SEQ ID 120.
XX
KW Human; colon; cytostatic; vaccine; colon cancer; colon disorder;
KW metastasis.

XX
OS Homo sapiens.
XX
PN WO200277234-A2.
XX
PD 03-OCT-2002.
XX
PF 31-OCT-2001; 2001WO-US48414.
XX
PR 31-OCT-2000; 2000US-244758P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Recipon H, Ghosh MG, Liu C;
XX
DR WPI; 2003-018928/01.
XX
PT New isolated colon-specific nucleic acid molecule, useful for treating
PT colon cancer, and diagnosing or monitoring the presence of metastases
PT of colon cancer in a patient -
XX
PS Claim 11; Page 192; 216pp; English.
XX
CC The present invention relates to human colon specific nucleic acids
CC (ABV93910-ABV94009) and proteins (ABP68360-ABP68435). The nucleic acids
CC and proteins are useful for treating colon cancer and colon disorders,
CC and diagnosing or monitoring the presence of colon disorders and
CC metastases of colon cancer in a patient.
XX
SQ Sequence 92 AA;

Query Match 33.8%; Score 514; DB 24; Length 92;
Best Local Similarity 100.0%; Pred. No. 1e-40;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 60

Qy         152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
               ||||||||||||||||||||||||||||||||
Db          61 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 92
```

Search completed: August 5, 2003, 13:08:23
Job time : 39 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 5, 2003, 13:07:04 ; Search time 14.5 Seconds
(without alignments)
805.365 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1520	100.0	276	4	US-09-094-964-2	Sequence 2, Appli
2	1520	100.0	348	3	US-08-652-265-6	Sequence 6, Appli
3	1520	100.0	348	3	US-08-834-497A-6	Sequence 6, Appli
4	1520	100.0	348	3	US-09-503-444A-6	Sequence 6, Appli
5	1513	99.5	276	4	US-09-094-964-1	Sequence 1, Appli
6	1513	99.5	348	3	US-08-652-265-2	Sequence 2, Appli
7	1513	99.5	348	3	US-08-834-497A-2	Sequence 2, Appli
8	1513	99.5	348	3	US-09-503-444A-2	Sequence 2, Appli
9	1513	99.5	348	4	US-09-277-457-2	Sequence 2, Appli
10	1513	99.5	348	4	US-09-679-729-2	Sequence 2, Appli
11	1509	99.3	348	3	US-08-652-265-8	Sequence 8, Appli
12	1509	99.3	348	3	US-08-834-497A-8	Sequence 8, Appli
13	1509	99.3	348	3	US-09-503-444A-8	Sequence 8, Appli
14	1502	98.8	348	3	US-08-652-265-4	Sequence 4, Appli
15	1502	98.8	348	3	US-08-834-497A-4	Sequence 4, Appli
16	1502	98.8	348	3	US-09-503-444A-4	Sequence 4, Appli
17	1493	98.2	276	4	US-09-094-964-3	Sequence 3, Appli
18	523	34.4	361	3	US-08-652-265-22	Sequence 22, Appl
19	523	34.4	361	3	US-08-834-497A-22	Sequence 22, Appl
20	523	34.4	361	3	US-09-503-444A-22	Sequence 22, Appl
21	517	34.0	364	4	US-08-914-372C-11	Sequence 11, Appl
22	514	33.8	365	3	US-08-652-265-23	Sequence 23, Appl
23	514	33.8	365	3	US-08-834-497A-23	Sequence 23, Appl
24	514	33.8	365	3	US-09-503-444A-23	Sequence 23, Appl
25	506	33.3	274	2	US-08-484-905-107	Sequence 107, App
26	506	33.3	274	3	US-08-481-985B-107	Sequence 107, App
27	506	33.3	274	3	US-08-370-476-107	Sequence 107, App
28	506	33.3	341	3	US-08-890-719-38	Sequence 38, Appl
29	505	33.2	365	2	US-08-484-905-97	Sequence 97, Appl
30	505	33.2	365	3	US-08-481-985B-97	Sequence 97, Appl
31	505	33.2	365	3	US-08-370-476-97	Sequence 97, Appl
32	504	33.2	274	2	US-08-484-905-108	Sequence 108, App
33	504	33.2	274	3	US-08-481-985B-108	Sequence 108, App
34	504	33.2	274	3	US-08-370-476-108	Sequence 108, App
35	504	33.2	365	2	US-08-484-905-100	Sequence 100, App
36	504	33.2	365	3	US-08-481-985B-100	Sequence 100, App
37	504	33.2	365	3	US-08-370-476-100	Sequence 100, App
38	503	33.1	274	1	US-08-222-851-1	Sequence 1, Appli
39	503	33.1	363	4	US-08-914-372C-37	Sequence 37, Appl
40	503	33.1	365	2	US-08-484-905-99	Sequence 99, Appl
41	503	33.1	365	3	US-08-481-985B-99	Sequence 99, Appl
42	503	33.1	365	3	US-08-370-476-99	Sequence 99, Appl
43	502	33.0	274	2	US-08-484-905-106	Sequence 106, App
44	502	33.0	274	3	US-08-481-985B-106	Sequence 106, App
45	502	33.0	274	3	US-08-370-476-106	Sequence 106, App

ALIGNMENTS

```
RESULT 1
US-09-094-964-2
; Sequence 2, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-2

  Query Match             100.0%;  Score 1520;  DB 4;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-142;
```

```
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-08-652-265-6
; Sequence 6, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
;    APPLICANT: Thomas, Winston J.
;    APPLICANT: Drayna, Dennis T.
;    APPLICANT: Feder, John N.
;    APPLICANT: Gnirke, Andreas
;    APPLICANT: Ruddy, David
;    APPLICANT: Tsuchihashi, Zenta
;    APPLICANT: Wolff, Roger K.
;    TITLE OF INVENTION: Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES: 44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Townsend and Townsend and Crew LLP
;      STREET: Two Embarcadero Center, Eighth Floor
;      CITY: San Francisco
;      STATE: California
;      COUNTRY: USA
;      ZIP: 94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/652,265
;      FILING DATE: 23-MAY-1996
;      CLASSIFICATION: 514
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Smith, William M.
;      REGISTRATION NUMBER: 30,223
;      REFERENCE/DOCKET NUMBER: 17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 576-0200
;      TELEFAX: (415) 576-0300
;  INFORMATION FOR SEQ ID NO: 6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 348 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
US-08-652-265-6

  Query Match             100.0%;  Score 1520;  DB 3;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 3
US-08-834-497A-6
; Sequence 6, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT: Thomas, Winston J.
;    APPLICANT: Drayna, Dennis T.
;    APPLICANT: Feder, John N.
;    APPLICANT: Gnirke, Andreas
;    APPLICANT: Ruddy, David
;    APPLICANT: Tsuchihashi, Zenta
;    APPLICANT: Wolff, Roger K.
;    TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES: 76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Pennie & Edmonds LLP
;      STREET: 1155 Avenue of the Americas
;      CITY: New York
;      STATE: New York
;      COUNTRY: USA
;      ZIP: 10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: Windows 95
;      SOFTWARE: FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/834,497A
;      FILING DATE: 04-APR-1997
;      CLASSIFICATION: 514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/652,265
;      FILING DATE: 23-MAY-1996
;      CLASSIFICATION: 514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/632,673
;      FILING DATE: 16-APR-1996
;      CLASSIFICATION: 514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/630,912
;      FILING DATE: 04-APR-1996
;      CLASSIFICATION: 514
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Poissant, Brian M.
;      REGISTRATION NUMBER: 28,462
;      REFERENCE/DOCKET NUMBER: 8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 650-493-4935
;      TELEFAX: 650-493-5556
;      TELEX: 66141 PENNIE
;  INFORMATION FOR SEQ ID NO: 6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 348 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-6

  Query Match             100.0%;  Score 1520;  DB 3;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
US-09-503-444A-6
; Sequence 6, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-6

  Query Match             100.0%;  Score 1520;  DB 3;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 5
US-09-094-964-1
; Sequence 1, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  276 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-094-964-1

  Query Match             99.5%;  Score 1513;  DB 4;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 6
US-08-652-265-2
; Sequence 2, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  348 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-652-265-2

  Query Match             99.5%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 7
US-08-834-497A-2
; Sequence 2, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-2

  Query Match             99.5%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 8
US-09-503-444A-2
; Sequence 2, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-2

  Query Match             99.5%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 9
US-09-277-457-2
; Sequence 2, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-277-457-2

  Query Match             99.5%;  Score 1513;  DB 4;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
```

```
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 10
US-09-679-729-2
; Sequence 2, Application US/09679729
; Patent No. 6509442
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 24065-004 DIV
;  CURRENT APPLICATION NUMBER: US/09/679,729
;  CURRENT FILING DATE:  2000-10-04
;  PRIOR APPLICATION NUMBER: 09/277,457
;  PRIOR FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-679-729-2

  Query Match           99.5%;  Score 1513;  DB 4;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 11
US-08-652-265-8
; Sequence 8, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-8

  Query Match           99.3%;  Score 1509;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.6e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 12
US-08-834-497A-8
; Sequence 8, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
```

```
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-8

  Query Match             99.3%;  Score 1509;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.6e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 13
US-09-503-444A-8
; Sequence 8, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-8

  Query Match             99.3%;  Score 1509;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.6e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 14
US-08-652-265-4
; Sequence 4, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-4

  Query Match             98.8%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-140;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 15
US-08-834-497A-4
; Sequence 4, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-4

  Query Match             98.8%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-140;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82
```

```
Qy          61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy         121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy         181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy         241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
               |||||||||||||||||| ||||||||||||||||
Db         263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

Search completed: August 5, 2003, 13:11:11
Job time : 15.5 secs